Edwards, T. L., & Challenor, P. (2013). Risk and uncertainty in hydrometeorological hazards. In J. Rougier, S. Sparks, & L. Hill (Eds.), Risk and Uncertainty Assessment for Natural Hazards. (pp. 100-150). Cambridge University Press, Cambridge.

Publisher's PDF, also known as Final Published Version

Link to publication record in Explore Bristol Research
PDF-document

5

Risk and uncertainty in hydrometeorological hazards

T. L. EDWARDS AND P. G. CHALLENOR

5.1 Introduction

Extreme weather and ocean hazards are inescapable. Weather is the state of the atmosphere, to which every part of the earth's surface is exposed. In coastal regions and at sea there are equivalent oceanic hazards, such as extreme waves. Every region has a local normal range for atmospheric and oceanic, or 'hydrometeorological', conditions, around which societal decisions such as habitation and infrastructure are made. But inevitably there are fluctuations outside these bounds that result in hazard events. In general it is too difficult and costly to be fully prepared for the rare extremes.

This chapter summarises the present-day risks and uncertainties associated with droughts, heat waves, extreme precipitation, wind storms and extreme ocean waves. How these may change under future climate change is discussed in Chapter 6. Other meteorological hazards, such as extreme cold, are not included here for reasons of space. Flooding and storm surge hazards are described separately in Chapter 7. Section 5.2 outlines the characteristics of the hazards. Section 5.3 discusses methods of assessing risks for hazards and their impacts. Section 5.4 describes organisations and tools of risk management and communication. Section 5.5 summarises.

5.2 Hydrometeorological hazards

This section gives an overview of the hazards (definition, triggers and scale) and their associated losses (exposure and vulnerability; types and quantification of loss).

5.2.1 Definition

The United Nations International Strategy for Disaster Reduction (http://www.unisdr.org) gives a general definition for a hydrometeorological hazard as a 'process or phenomenon of atmospheric, hydrological or oceanographic nature that may cause loss of life, injury or other health impacts, property damage, loss of livelihoods and services, social and economic disruption, or environmental damage'. Most hydrometeorological hazards are true 'extremes' in the sense that they are not distinct events but a consequence of entering the high or low end of local climatic variation (e.g. IPCC, 2011). Exceptions to this are cyclonic storms (such as hurricanes), which are discrete atmospheric patterns of circulating wind.

Hydrometeorological hazards thus do not have triggers or inception events, though particular atmospheric or oceanic states make them more likely (Section 5.2.2). These 'extremal' characteristics lead naturally to probabilistic risk assessments, quantified in terms of the probability of being in a particular region of climate state space (Chapter 6). The term 'risk' is defined in Chapter 2.

The definition of a hydrometeorological hazard varies by location and by the type of risk assessment. Extremes must be defined relative to regional climatology (mean climate) rather than fixed thresholds, because local acclimatisation and infrastructure affect the seriousness of the impacts. And for risk assessments, definitions range from simple climate-based metrics such as daily maxima, which are used for quantifying the probability of occurrence, to more complex multivariate definitions, for describing impacts on health or infrastructure.

Extreme events are typically defined in terms of threshold exceedance, where thresholds are defined relative to local climatology (such as the 90th percentile of daily maximum temperatures at a given location for a particular time period) so as to quantify events of a fixed rarity. Alternatively, an absolute threshold may be used for all locations so as to quantify events of a fixed intensity. Duration of hazard is also important, because the impacts of extreme events are greater when extreme conditions prevail over extended time periods. The joint CCI/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) recommended 27 indices of extreme temperature and precipitation so as to standardise comparisons globally (Alexander *et al.*, 2006). These have been widely used, in particular for the most recent assessment report by the Intergovernmental Panel on Climate Change (IPCC, 2007a).

The path into the extreme also affects the hazard. A slow increase in dryness before a drought may result in a more severe event than a sudden transition, due to the cumulative (integrated) impacts, while a heat wave at the end of the summer typically results in lower mortality rates than one at the start due to acclimatisation during the season and the 'harvesting' of weak individuals (Baccini *et al.*, 2008). This context is not always considered in risk assessments, partly because sub-dividing extreme data into categories exacerbates any problems of low numbers of observations.

Once a hazard is defined, there are two types of risk assessment: physical model-based forecasting (Section 5.3.2), which is the prediction of specific hazard events in the near future, and statistical analysis (Section 5.3.3), which is the estimation of frequency over the longer term. For the latter, two very commonly used metrics are the 'return period', which is the inverse of the probability of an event with specified magnitude at a given location occurring during the next year, and equivalently the 'return level', the magnitude associated with a given return period. A more precise characterisation is given in Chapter 2. Return periods are defined for a particular hazard magnitude and location, but there are other relevant characteristics such as duration, spatial extent and (in the case of storms) path. The return period may depend on how these aspects are combined (Della-Marta *et al.*, 2009). A return period is a statement of probability at a given time, not a projection for the future; if the data are stationary and independent, the return period can be interpreted as the expected waiting time until the next event. Return periods are usually estimated from the

observational record, but are also estimated from re-analysis datasets (observations assimilated into a physical weather model simulation) such as ERA-40 (Uppala *et al.*, 2005) or from stand-alone simulations (Chapter 6).

Each hydrometeorological hazard has a typical definition, or in some cases several, according to the event type and the intended use of the information. These definitions are outlined below.

5.2.1.1 Extreme precipitation

Extreme precipitation events are usually defined simply by their magnitude, as an exceedance of a climatological threshold such as the 99th percentile (e.g. Barnett *et al.*, 2006). The definition may include a measure of persistence such as a five-day total (Beniston *et al.*, 2007). Frequency and ratio of extreme precipitation relative to local climatology are also important due to risk of inundating flood control systems (Chapter 7). Four indices of precipitation are shown in Figure 5.1.

5.2.1.2 Heat waves

Heat waves are persistent periods of high temperature. The persistence is significant for human health risk, particularly for individuals that are not usually vulnerable (e.g. Joacim *et al.*, 2010). Heat wave definitions always vary by region, because humans acclimatise to their local temperature and humidity. The simplest are expressed in terms of the duration of the daily maximum temperature, T_{max}. One commonly used definition is T_{max} exceeding climatology by 5°C for at least five consecutive days, either any such period (Frich *et al.*, 2002) or the longest such period in the year (Tebaldi *et al.*, 2006). A more complex definition
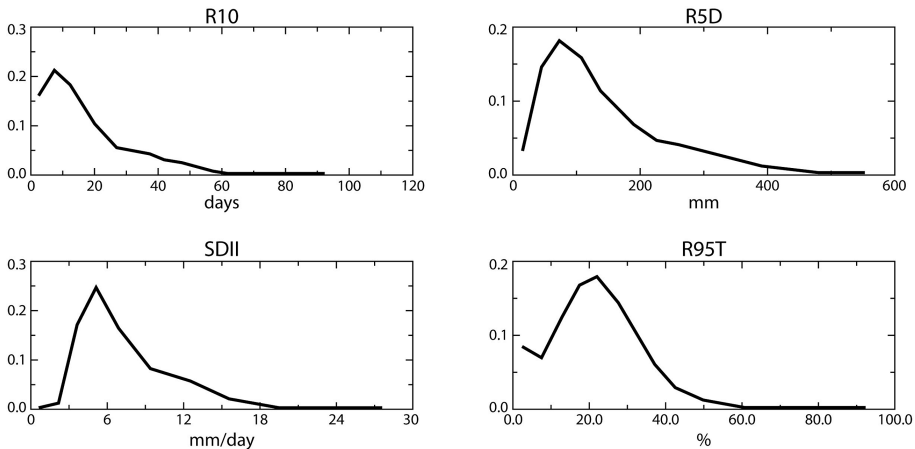
Figure 5.1 R10: number of days with precipitation $\geq 10\,mm\,d^{-1}$; SDII: simple daily intensity index: annual total / number of days with precipitation $\geq 1\,mm\,d^{-1}$; R5D: maximum five-day precipitation total; and R95T: fraction of annual total precipitation due to events exceeding the 1961–1990 95th percentile. Adapted from Frich *et al.* (2002) © Inter-Research.

using a critical threshold ($T_c = 25°C$ for Central Europe) is used by Kyselý (2009): a continuous period during which mean T_{max} exceeds T_c by 5°C in at least three days, the mean T_{max} for the whole period exceeds T_c by 5°C, and T_{max} does not fall below T_c. Temperature variability is much smaller in some regions than others, so fixed-excess definitions such as these are not appropriate for global studies: instead the excess should be set relative to local climatology (Alexander *et al*., 2006).

Duration of high T_{max} is generally thought to be insufficient information for assessing human health risks. Persistently high night-time temperatures prevent core body temperatures from cooling, and high humidity levels generally reduce tolerance to heat; Gosling *et al*. (2009) discuss these issues in more detail. So warning systems increasingly use sophisticated metrics more suitable for human health. The Heat Index of the National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) combines temperature and humidity criteria (Robinson, 2001). This is a measure of 'how hot it feels': for example, an air temperature of 96°F (36°C) and relative humidity of 65% results in a Heat Index of 121°F (49°C). 'Synoptic' definitions of heat waves take into account the overall type of meteorological conditions and the previous impact of these conditions on mortality and health. One example used widely in the United States and in parts of Europe is the Spatial Synoptic Classification (SSC; Sheridan, 2002; Sheridan and Kalkstein, 2004). This uses several meteorological variables (temperature, dew point, cloud cover, pressure and wind) to classify weather into one of ten types and quantifies the persistence, relative timing within the summer season and particular meteorological characteristics of that type. Souch and Grimmond (2004) and Smoyer-Tomic *et al*. (2003) describe other heat wave and heat stress indices. Some studies define a heat wave according to the characteristics of a particular event, such as the 2003 European heat wave, to estimate the return period (Schär *et al*., 2004) and changing risk of such events in the future (Beniston, 2004). However, the number of analogues found is sensitive to the definition of the event: whether in terms of absolute temperature or threshold exceedance, and whether individual or consecutive days (Beniston, 2004).

5.2.1.3 Droughts

Droughts are persistent dry periods that last from a few days to several years. There are various types (IPCC, 2007b; Burke and Brown, 2008), including meteorological drought (low precipitation), hydrological drought (low water reserves), agricultural drought (low soil moisture) and environmental drought (a combination of these). Droughts are quantified using one of various indices. Some are based solely on precipitation, such as the Standard Precipitation Index (SPI) and the China-Z Index (CZI; Wu *et al*., 2001), while others take into account additional factors such as temperature, potential evaporation, runoff and soil moisture. A commonly used metric of 'cumulative departure of moisture supply' from the normal is the Palmer Drought Severity Index (PDSI: Palmer, 1965). This is based on precipitation, temperature and soil moisture, though has a number of different formulations (Robeson, 2008; Dai, 2011). In fact, more than 150 definitions of drought exist

(Nicolosi *et al.*, 2009). These have different strengths and weaknesses, and the most suitable choice depends on the problem, context and user (Kallis, 2008).

Drought severity is described using thresholds of an index such as the first (extreme), fifth (severe) or twentieth (moderate) percentiles of present-day climate. Different thresholds may be set for each region to account not only for local climatology but also environmental conditions that increase vulnerability. The PDSI is calibrated for the United States and has categories −1.0 to −1.9 for mild drought, −2 to −3 for moderate drought, −3 to −4 for severe drought, and beyond −4 for an extreme drought (Leary *et al.*, 2008).

The rate of onset and cessation of drought are also important. A slow transition to drought may allow more time for mitigation, but could also worsen the impacts due to the cumulative effects before and after the defined drought phase.

5.2.1.4 Wind storms

Wind storms include hurricanes, typhoons, tornados, thunderstorms and extratropical cyclones. Tropical wind storms are 'warm core' cyclones, generated by convection of warm air over warm sea surface waters, while mid-latitude storms are 'cold core' and form along weather fronts. Wind storms are usually defined in terms of wind speed, with an absolute or relative threshold. A hurricane is a tropical storm with wind speeds of 74 mph (33 m s^{-1}, Category 1 on the Saffir–Simpson scale, or about 12 on the Beaufort scale). 'Storm force' winds in the UK have mean speed greater than 55 mph (25 m s^{-1}, 10 on the Beaufort scale) and gusts greater than 85 mph (38 m s^{-1}) affecting most of the country for at least six hours (UK Government Cabinet Office, 'UK Resilience': available from http://www.cabinetoffice.gov.uk). Della-Marta *et al.* (2009) compare five wind-based indices for storm magnitude and spatial extent, finding some are more suitable for quantifying strength relative to climatology and others for absolute strength.

Wind speed is not the only important metric for wind storms. The destruction caused by Hurricane Katrina in the United States in 2005 (Section 5.2.4.3) was due not to high wind speed but slow storm transit speed, heavy rainfall and a massive associated storm surge, as well as the vulnerability of the point at which it struck the coast. Storms are therefore often characterised with additional quantities such as sea-level pressure, vorticity, precipitation, lightning, storm surges or accumulated cyclone energy (ACE) (e.g. Beniston *et al.*, 2007).

Mid-latitude wind storms, or extratropical cyclones, in the Northern Hemisphere occur at a range of spatial scales. They are often studied at 'synoptic' (large) scales, 100–1000 km spatially and 6–72 hours in duration. Extreme synoptic storms have relatively moderate wind speeds compared with hurricanes or tornados, but cause great wind damage and economic loss due to the integrated effects of their large spatial extent. Smaller-scale structures may also develop within a large-scale storm. 'Sting-jets' (Browning, 2004) are characterised by very intense wind gusts and scales of order of tens of kilometres. 'Bombs' are rapidly deepening extratropical cyclones (Sanders and Gyakum, 1980) that develop in isolation from other vortices and are often characterised by spatial scales of around 100 km. Activities at these different spatial scales can interact and develop two-way feedbacks that may lead to explosive intensification and cyclogenesis (Rivière and Orlanski, 2007; Fink *et al.*, 2009).

5.2.1.5 Extreme waves and currents

Wind storms over the ocean generate large waves whose height depends on the wind speed, duration and fetch (distance over which the wind blows), and submarine earthquakes cause very large (tsunami) waves, but not all large waves are associated with storms and earthquakes. 'Freak waves' generated by nonlinearities in the wave field itself can be a significant hazard. Freak waves can also be generated when energy is focused by local currents and sea bed topography ('spatial focusing'). The Agulhas current along the east coast of South Africa, for example, meets swell from the Southern Ocean and creates 'holes in the sea' that cause severe damage to ships and loss of life. Freak waves are also created by nonlinear wave–wave interactions ('nonlinear focusing': Tucker and Pitt, 2001; Dysthe *et al.*, 2008). A further proposed mechanism is simple linear superposition of waves ('dispersive focusing': Kharif and Pelinovsky, 2003).

The wave field is regarded as a stationary random process over short periods of time (of order 30 minutes). This stochastic process is characterised by a frequency spectrum as a function of the direction of travel of the waves. Summary statistics are then derived from these spectra, such as 'significant wave height', which is four times the square root of the area under the spectrum (the total energy of the wave field). Extreme wave conditions are characterised in terms of the summary statistics: for example, freak waves are often defined as at least twice the significant wave height, which for linear waves is the expected height of the highest wave in a three-hour period (Dysthe *et al.*, 2008). An overview is given by Tucker and Pitt (2001). Although waves are the most important environmental force on offshore structures such as oil rigs, ocean currents can also be a serious hazard.

5.2.2 Triggers

Hydrometeorological hazards have no distinct trigger, but they can often be linked to certain states of the climate. Natural variability in the form of fluctuations and oscillations is present in the climate at all timescales, from hours to centuries and longer (Section 5.3.1), and the phases of this variability affect the likelihood of extreme events. If these are identified then particular climate states can be useful indicators of hazard risk in other regions or in the near future. One is the El Niño Southern Oscillation (ENSO), a measure of sea surface temperature anomalies in the tropical eastern Pacific, which affects the likelihood and location of extreme warm events, extreme precipitation and tropical cyclones (e.g. Rusticucci and Vargas, 2002; Camargo *et al.*, 2007). Another is the North Atlantic Oscillation (NAO), usually defined as the difference in the normalised surface air pressure between Iceland and the Azores or the Iberian peninsular. The NAO varies on decadal timescales and is a good proxy for the North Atlantic storm track. It has been linked with Atlantic cyclone intensity and with winter storms, precipitation and wave conditions in the eastern North Atlantic and Europe (e.g. Woolf *et al.*, 2002; Elsner, 2003; Mailier *et al.*, 2006; Vitolo *et al.*, 2009). Another example of a particular climate state is 'blocking', where an atmospheric pattern persists and prevents normal local weather from moving through the area, which results in

very stable weather. If this happens to occur at a time of high temperature or precipitation, a period of extremes can last for days or weeks.

There are physical connections between the different hazards. The 'hot and dry' hazards (heat waves, drought) are linked to each other through meteorological processes such as evapotranspiration and atmospheric convection, while the 'wet and windy' hazards (precipitation, storms) are linked through low atmospheric pressure. The hydrological cycle and water conservation connect droughts with extreme precipitation (Held and Soden, 2006). Hydrometeorological hazards also trigger other types of hazard: drought and heat waves cause wildfires; persistent extreme precipitation causes flooding and landslides; and wind storms cause storm surges.

5.2.3 Scale

Different parts of the world experience hazards of different scales – i.e. severity and frequency – due to the dependence of local climatic variability on local atmospheric patterns (e.g. circulation belts, jet streams, monsoons), topography (elevation and proximity to mountains and coast) and land and water use (e.g. intensive agriculture, deforestation, water consumption and redirection).

The scale of 'wet and windy' hydrometeorological hazards is mostly determined by natural factors. Extreme precipitation events are most severe in the mid-latitudes, in monsoon regions and on the windward side of mountains, while storm events are most severe in the tropics and in open areas such as coasts. One human factor, aerosol pollution, affects precipitation by acting as cloud condensation nuclei, but the interactions are complex: heavy rain events may increase at the expense of light rain events, and the net suppression or enhancement depends on particle size (Levin and Cotton, 2008; Li *et al.*, 2011). In an average year, ten tropical storms develop in the Gulf of Mexico, Caribbean Sea or Atlantic Ocean, of which six develop into hurricanes (Townsend, 2006). Mid-latitude winter storms are the most frequent natural disaster for Central Europe, while sting-jets are also believed to be a major cause of wind damage in northern Europe (Fink *et al.*, 2009). In general the highest storm-related waves are in the North Atlantic in winter (Holliday *et al.*, 2006). Waves in the Southern Ocean are smaller but with little seasonal variability, so the Southern Ocean is always rough, whereas the North Atlantic and North Pacific are relatively calm during the summer months (e.g. Young, 1999; Woolf *et al.*, 2002).

The scale of 'hot and dry' hydrometeorological hazards is influenced by human and natural factors. Heat waves are most severe in areas of inland desert, semi-desert and with Mediterranean-type climates (i.e. tropical and subtropical regions), and are often associated with an unusual location of the jet stream. Urban areas experience higher temperatures than rural areas due to the heat island effect. Drought hazards are most severe in areas with low precipitation, strong winds, high temperature, vulnerable soil types (such as high sand content) and low water reserves. These regions are often subtropical, mid-continental or

primarily supplied by glacial meltwater. Humans amplify droughts through deforestation, intensive farming and high population densities (IPCC, 2011), which leave soil exposed, increase runoff and deplete water reservoirs such as aquifers. An example of this is the Anglian region in eastern England, where low water availability is exacerbated by very high demand from agriculture. Areas that have suffered from severe drought include north-eastern and western Africa, northern and central China, western and mid-continental north America, the Amazon basin and Australia (e.g. Dai, 2011).

Regional variation in the magnitude of extremes may be illustrated with some examples. In the UK, the highest recorded daily rainfall during one of the most severe flooding events in living memory (November 2009) was 316 mm (UK Met Office website: http://www. metoffice.gov.uk). But the current world record for a 24-hour period is nearly six times greater, 1.825 metres, recorded in January 1966 by the island Réunion, east of Madagascar (World Meteorological Organization: http://wmo.asu.edu). The most significant UK wind storms in recent decades were in 1987, with a mean wind speed of 50 mph and gusts up to 115 mph, and in 1990, with mean wind speed of 69–75 mph and gusts up to 104 mph (UK Government Cabinet Office, 'UK Resilience': available from http://www.cabinetoffice.gov. uk). Hurricane Katrina in 2005 was classified as Category 5 (greater than 155 mph) but weakened to Category 3 (111–130 mph) before reaching land in Louisiana. The highest directly measured gust speed is 253 mph, during a tropical cyclone in April 1996 on Barrow Island off the north-west coast of Australia, and the highest recorded speed (inferred using radar) is 302 mph in May 1999 in Oklahoma (World Meteorological Organization: http:// wmo.asu.edu). The highest wave height measurements are around 30 m (Kharif *et al.*, 2009), though such high values are not necessarily classified as 'freak' according to the usual definition of twice the significant wave height (Section 5.2.1.5). During Hurricane Ivan over the Gulf of Mexico in 2005, for example, the maximum wave height was 26.3 m, while twice the highest significant wave height was 30.8 m (Forristall, 2005). During the UK heat wave of August 1990, temperatures of 32°C or more (a critical threshold: Section 5.4.1.2) were recorded in virtually all parts of England and some parts of Wales, 11.5°C hotter than the 1971–2000 average for August maximum temperatures (UK Met Office website: http://www.metoffice.gov.uk); the most persistent UK event was in 1976, with 32°C exceeded for 15 consecutive days (UK Government Cabinet Office, 2008). During the 2010 Northern Hemisphere heat wave, a temperature of 38.2°C was recorded in Moscow (Barriopedro *et al.*, 2011). The number of European extreme drought events during the twentieth century, with PDSI less than or equal to −4, is shown in Figure 5.2 (Lloyd-Hughes and Saunders, 2002).

5.2.4 Loss

This section gives an overview of the losses caused by hydrometeorological hazards, outlining the entities that are exposed and vulnerable, the types of loss that may occur and the quantification of these losses.
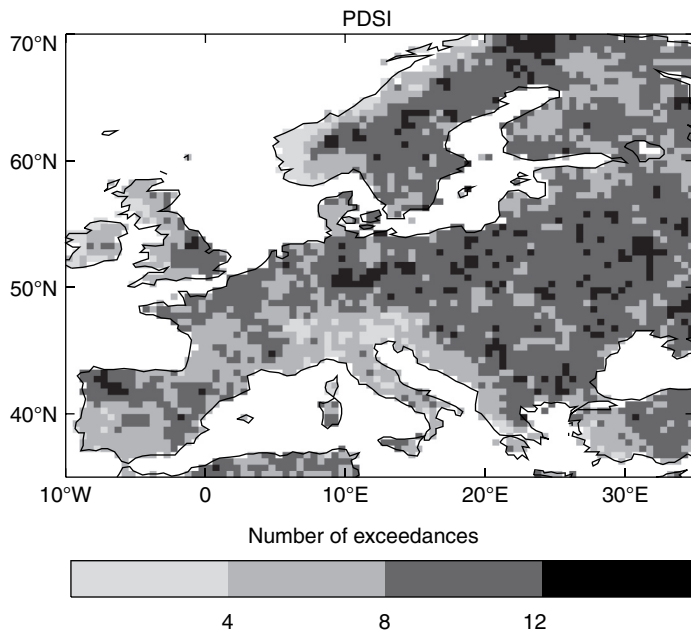
Figure 5.2 Number of extreme drought events (PDSI less than or equal to −4) by grid cell for Europe during the period 1901–1999. (Lloyd-Hughes and Saunders, 2002) © Wiley.

5.2.4.1 *Exposure and vulnerability*

Meteorological hazards have global reach: they affect everything under the sky. Hydrological hazards are almost as pervasive, because exposure comes from proximity to bodies of water (e.g. waves) and from limited water supply (droughts). Some types of hazard do occur mainly in one region, such as hurricanes in the tropical Atlantic or typhoons in the Pacific.

The entities exposed to hydrometeorological hazards are therefore wide-ranging: humans; agriculture (including crops, livestock and irrigation); infrastructure (including transport, power and housing); and the environment (including forests, surface soil, ecosystems and habitats) (e.g. Beniston *et al.*, 2007; UK Environment Agency, 2009; UK Government Cabinet Office, 2008). Vulnerability varies by location: poverty and inefficient infrastructure reduce the ability to prepare for hazards and to recover (IPCC, 2011); local architectural styles, construction practices, building regulations and enforcement affect the vulnerability of buildings (RMS, 2005); and vulnerability is greater if the climate is close to habitable limits, though prior conditioning to warm climate or recent high temperatures reduces vulnerability to heat waves.

Densely populated areas are particularly vulnerable, because not only are more people exposed but the effects are amplified: low water reserves in the case of drought, and

mortality or injury due to damaged or collapsing buildings in the case of extreme precipitation and wind storms. Nearly 50% of the world's most populated areas are highly vulnerable to drought (USDA, 1994). However, rural areas are also vulnerable: for example, trees in densely wooded areas are felled by strong winds, causing damage and injury.

Social and cultural practices are also important, particularly for heat waves. The most vulnerable individuals are those living alone, ill, bedridden, housebound, elderly, over-weight, lacking nearby social contacts, lacking access to transportation or lacking sufficient domestic cooling (Klinenberg, 2002). Air pollution also exacerbates the health risk (Souch and Grimmond, 2004). The European heat wave in 2003 and the two Chicago heat waves during the 1990s mostly affected elderly people in urban areas (United Nations, 2006).

5.2.4.2 Types of loss

Humans are killed, injured or made ill by every type of hydrometeorological hazard: wind storms and waves cause physical injury; drought causes dehydration, related diseases and malnutrition due to famine; heat waves cause dehydration, heat stroke, oedema (swelling), rashes and cramps (reviewed by Kovats and Hajat, 2008); and extreme precipitation can lead to waterborne diseases (Curriero *et al*., 2001) and, in combination with high temperatures, to malaria (Githeko and Ndegwa, 2001). But the impacts on humans are not limited to mortality and morbidity: sustained drought can cause mass migration, political unrest and violent conflict over natural resources (Gleick *et al*., 2006), while heat waves decrease productivity and increase violent crime (Smoyer-Tomic *et al*., 2003). Drought, heat waves and extreme precipitation reduce agricultural yields through high temperature, low water availability and soil erosion. Extended drought can lead to famine, amplifying problems of health and social stability (Gleick *et al*., 2006).

Infrastructure, property and business are also harmed. Energy structures (including power stations, oil and gas platforms, refineries, pipelines and renewable energy devices) are damaged and power supplies interrupted or reduced by wind storms and extreme waves (physical damage), extreme precipitation (flooding), heat waves (overheating) or drought (overheating; reduced flow through hydroelectric dams). Transport and telecommunication links are damaged by wind storms, rain and (in the case of ships) extreme waves, and during heat waves road surfaces are susceptible to melting, vehicle engines to overheating and railway lines to buckling (De Bono *et al*., 2004). Water supplies are affected by extreme precipitation (polluting), drought (decreasing sources) and heat waves (buckling pipes).

Hydrometeorological hazards may be natural, but can severely harm the natural environment. Drought and extreme precipitation cause habitat damage and soil erosion, and the latter can lead to desertification and dust storms. Wind storms also harm ecosystems (particularly forests, e.g. Bründl and Rickli, 2002). Wildfires are caused by drought and heat waves. Heat waves may affect glaciers, permafrost and rock stability (De Bono *et al*., 2004). Water pollution from extreme precipitation has negative impacts on ecosystems as well as human health.

5.2.4.3 Quantification of loss

Quantification of loss is more straightforward for some types of harm and damage than others. The most often used metrics relate to human health and financial cost.

Mortality can be expressed in absolute numbers, though is more usefully expressed as 'excess death rate', the difference between reported and typical deaths for a given time period. But it can be difficult to quantify the effect of 'harvesting', the phenomenon by which the death of vulnerable individuals is hastened but overall mortality is not increased (e.g. Baccini *et al.*, 2008). Ill health can be quantified with hospital records, but these data are less available and less easily interpreted than mortality data (Sheridan and Kalkstein, 2004). Mortality rates in the United States are highest from extreme heat and hurricane hazards (Figure 5.3), though the latter are dominated by Hurricane Katrina, which caused over 1000 deaths. Negative impacts of heat extremes rose during the period 1995–2004, especially in Europe (United Nations, 2006), where the 2003 heat wave led to the loss of more than 70 000 lives (Robine *et al.*, 2008). Waves contributed to one of the most severe environmental disasters to occur in the UK, when flood defences on the east coast of England were breached in 1953 by a combination of high tides, storm surge and large waves: 307 people were killed in the area (UK Government Cabinet Office, 2008).
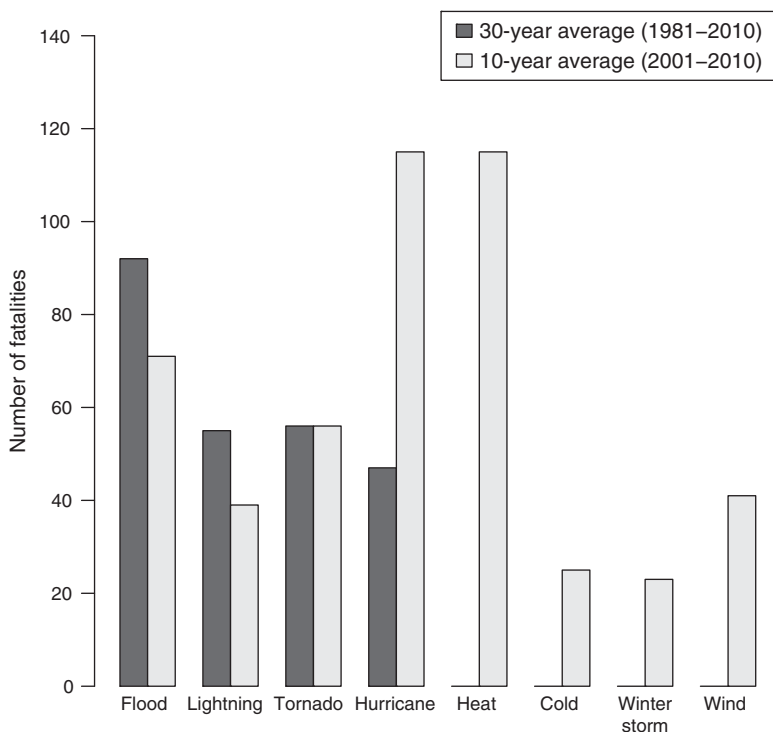

Figure 5.3 Mean annual weather fatalities for the United States, Puerto Rico, Guam and the Virgin Islands. Data from the National Weather Service.

Financial losses due to physical damage of property and infrastructure or interruption of services are typically quantified by insurance claims or economic effect. The former is not comprehensive because insurance coverage is rare in many parts of the world. Both insured and uninsured losses from natural hazards have risen during the twentieth century due to increases in exposure (population; habitation of susceptible areas such as coasts) and vulnerability (asset ownership; changes in social structure, infrastructure and environmental conditions) (Grossi and Kunreuther, 2005; Dlugolecki *et al.*, 2009). Climatic extremes caused more than 170 'billion-dollar events' (in insurance terms) globally during the second half of the twentieth century, compared with 71 due to earthquakes (Beniston *et al.*, 2007). The most important were tropical cyclones and mid-latitude winter storms (Fink *et al.*, 2009), followed by floods, droughts and heat waves. Hurricane Katrina is the most destructive natural disaster in the history of the United States (Townsend, 2006). It caused destruction or severe damage to around 300 000 homes; at least ten oil spills; an initial shutdown of more than 91% of oil and 83% of gas production in the Gulf of Mexico region (RMS, 2005); and estimated damages of US$96 billion (Townsend, 2006). In December 1999, the three storms Anatol, Lothar and Martin caused a total insured loss of more than €10 billion across Europe (Fink *et al.*, 2009) and in 2007 another, Kyrill, caused €4–7 billion of insured losses and the loss of electricity for at least two million homes (Fink *et al.*, 2009). Lawton (2001) reports an estimate that freak waves were responsible for 22 of the 60 losses of super-carriers (very large aircraft carriers) to sudden flooding between 1969 and 1994 (Kharif and Pelinovsky, 2003).

Harm to ecosystems and environmental services is harder to quantify, though its importance is increasingly recognised (e.g. the Millennium Ecosystem Assessment; Hassan *et al.*, 2005). Harm may be expressed in terms such as loss of biodiversity, habitat or natural resources. Mid-latitude storms have caused significant environmental losses: in 2007, storm Kyrill uprooted an estimated 62 million trees across Europe (Fink *et al.*, 2009).

5.3 Risk assessment

Risk assessment for hydrometeorological hazards has a very long history. Humans have attempted to forecast weather from patterns on the earth and in the stars since antiquity, recorded its extremes in historical documents and measured its variations since the invention of each meteorological instrument (e.g. thermometer, rain gauge, barometer). The longest instrumental record of temperatures is the Central England Temperature (CET), which began in 1659 (Parker *et al.*, 1992). The field of numerical weather prediction, proposing equations to describe the physical behaviour of the atmosphere, has existed since the turn of the twentieth century, and computers have been used to solve these equations since the 1950s (Lynch, 2008). Despite this long history, hydrometeorological hazards, and their impacts, are still challenging to predict (Section 5.3.1).

The word 'risk' is very often applied in this field, as in many others, only to the likelihood of a hazard occurring rather than in the full sense incorporating damage and loss; an

exception to this is catastrophe modelling (Section 5.3.3). There are two types of risk assessment, based on physical modelling and statistical analysis. Physical models, also known as 'dynamical models' or 'simulators', are mathematical equations written in computer code to simulate the behaviour of the atmosphere, ocean and land. They are used to forecast specific hazard events – their magnitude, extent and timing – in order to issue warnings or take steps to mitigate the hazard (Section 5.3.2). Statistical analysis of hazard events (Section 5.3.3) assesses the long-term characteristics of a particular hazard type, expressed in terms of return period or dependence on other variables. Some risk assessments include both physical and statistical elements, in particular hurricane forecasts (Sections 5.3.2, 5.3.3.3).

5.3.1 Challenges in risk assessment

The short length of the observational records limits estimation of return periods for rare hydrometeorological events. For some variables the record length is 150 years or more, but in general observations or reliable re-analysis data are only available for a few decades. Hydrometeorological hazards also have their own particular challenges: the earth system's internal variability, and non-stationarity due to the influence of forcings. These are discussed below.

5.3.1.1 Internal variability

The earth's atmosphere and oceans are continuously changing due to the influence of natural and anthropogenic 'forcings' that produce a radiative imbalance at the top of the atmosphere (Section 5.3.1.2). But even if there were no imbalance, and the earth system were at equilibrium, there would still be fluctuations of unforced 'internal' variability. The peaks and troughs of these fluctuations give rise to extreme events. Internal variability derives from two sources: chaos and oscillations. These are challenging to simulate (Section 5.3.2) and must be accounted for in statistical analyses (Section 5.3.3).

Chaos arises from the sensitivity of the atmosphere's evolution to small changes in the 'initial conditions', a consequence of the earth system's complex and nonlinear nature. The atmosphere and ocean are turbulent, with nonlinear feedbacks involving exchanges of heat and water, chemical reactions, and so on, between every part of the air, sea, land, ice and life. Nonlinearity and complexity lead to chaos, but chaos does not imply randomness. A chaotic system can be deterministic: in such a system, perfect knowledge of the initial conditions and response would enable prediction over long timescales. But knowledge of the earth system cannot be perfect. Observations have finite coverage and are uncertain. Physical models have finite temporal and spatial resolution and are imperfect, because many processes are unknown or difficult to represent, and the observations available to evaluate them are limited. So chaos, combined with imperfect knowledge, results in 'limited predictability', where uncertainties in initial conditions are greatly amplified in predictions beyond a few days. The limit of predictability is thought to be about two weeks (Lorenz, 1965).

Limited predictability is equivalent, in practice, to a random element, so most forecasting models have stochastic components (Section 5.3.2).

Oscillations in the state of the atmosphere and ocean move around the earth at all timescales, from weeks (such the Madden–Julian Oscillation, MJO) to years (such as the ENSO and NAO; Section 5.2.2) and longer. These oscillations are difficult to simulate realistically in physical models because they must emerge spontaneously from the smaller-scale behaviour described by the equations; some physical models are more successful at this than others (IPCC, 2007a). Oscillations are also a challenge in statistical analysis of observations, because they are difficult to separate from long-term trends. But these longer-term changes do bring an important advantage: recent improvements in understanding and measuring their persistent influence on atmospheric conditions allow forecasts to be made of spatial and temporal averages (such as the number of hurricanes in the coming year) with much longer lead times than Lorenz's fortnight, albeit with larger uncertainties (Sections 5.3.2, 5.3.3.3).

5.3.1.2 Non-stationarity

The atmosphere and oceans are continuously 'pushed' towards warmer or cooler states by a radiative imbalance at the top of the atmosphere. This imbalance is the net result of 'forcings', both natural and anthropogenic in origin, which affect the amount of solar radiation that warms the planet or escapes to space. During the instrumental record (the past 150 years or so), the most important forcings are: greenhouse gases and tropospheric ozone that warm the planet (increasing due to industrial emissions and release from natural reservoirs such as forests and wetlands, and varying due to natural cycles); incoming solar radiation (varying in regular cycles and long-term trends); and sulphate aerosols that have a cooling effect (varying due to industrial emissions and volcanic eruptions). More details are given in Chapter 6. Over the twentieth century, the net forcing has resulted in an increase of about 0.7°C in global mean temperature (IPCC, 2007a), with larger changes over land and in regions with strong positive feedbacks such as the Arctic. The twentieth-century response is relatively small compared with internal variability such as ENSO, but reflects an increase in energy in the system that is thought to have shifted the statistical distribution of several aspects of weather (IPCC, 2011). The impacts of hazard events are also non-stationary, due to changing social and cultural factors: for example, mortality due to heat waves is decreasing in the United States due to increased use of air conditioning, improved healthcare and improved public awareness of impacts (Souch and Grimmond, 2004). As well as interannual non-stationarity, hydrometeorological hazards and their impacts vary with the seasonal cycle. The influence of these oscillations, like those of internal variability in the previous section, aids forecasting on seasonal timescales.

Non-stationarity should therefore be accounted for, where present, in statistical analysis of observational records. Furthermore, warming is expected to continue over the next few decades and longer, due to the slow response of the oceans to past forcing, the long lifetime of greenhouse gases, the cumulative effect of a (likely) slow rate of reducing greenhouse gas emissions, and a (likely) reduction in sulphate aerosols from pollution.

Hydrometeorological hazards are part of the climatic spectrum and likely to be affected by this warming: estimates of return periods, which inherently assume stationarity, have limited applicability for these hazards in the future. Estimating risk of hydrometeorological hazards under future climate change is discussed in Chapter 6.

5.3.1.3 Hazard- and impact-specific challenges

Some hydrometeorological hazards, or aspects of hazards, are more challenging to simulate or analyse than others. Temperature is the most straightforward to simulate because the dominant drivers of regional temperature patterns are latitudinal insolation patterns and the distribution of land around the globe, which are well known (IPCC, 2007a). Precipitation patterns are influenced by these and also by small-scale processes such as atmospheric instabilities, the flow of air over local topography and very small-scale processes such as formation of cloud condensation nuclei (IPCC, 2007a); these are difficult to simulate (Stephens *et al.*, 2010) and in general must be represented in a simplified form (Section 5.3.2). Precipitation also responds nonlinearly to forcings, which complicates statistical analysis. Forecasting hurricane tracks is improving over time, but forecasting their intensities and landfall is much less successful (Camargo *et al.*, 2007). Droughts and heat waves are very influenced by soil moisture, but a global observational network is currently at an early stage (Dorigo *et al.*, 2011), so this important driver must be inferred from precipitation deficits for hazard assessments (Hirschi *et al.*, 2011). However, the global extent of hydrometeorological hazards may be of some advantage: research techniques of developed nations can be transferred to regions with fewer resources, and larger datasets can aid statistical analysis.

Assessing impacts and loss is also more challenging for some vulnerable entities than others: human mortality and health impacts, for example, depend on complex interactions between many factors that are hard to predict, such as societal decisions (CCSP, 2008), and are quantified differently depending on the relative importance of individual behaviour (sociology and environmental psychology), financial security (economics) and national capability and decision-making (political science) (Alcamo *et al.*, 2008).

5.3.2 Physical modelling

This section describes physical modelling of hydrometeorological hazards and impacts and quantification of uncertainty in their predictions. Physically based, dynamical models – simulators that attempt to represent real-world processes – are essential tools in forecasting the specific evolution of a state through time. They can also, if run multiple times simulta-neously (described below) or for a long time period (Chapter 6), be used to assess the probability distribution of a future hazard or impact. The advantage of using a physically based model to assess event probabilities, rather than a statistical analysis based on the historical record, is that it can account for non-stationarity, because the boundary conditions that drive its evolution (discussed below) can be varied in time and space. The disadvantage

is that it is hugely expensive, infeasibly so for events that have return periods of several years. Therefore assigning probabilities to low-probability, high-magnitude hazards tends to be done using statistical analysis (Section 5.3.3). Statistical models are also used for forecasting if there are temporally persistent states (such as ENSO) that influence the hazard (such as hurricane intensity), by assessing the future probability distribution given the specific present conditions.

5.3.2.1 Types of physical model

Numerical weather prediction (NWP) models are complex, computationally expensive simulators of the atmosphere and ocean that are used to forecast day-to-day weather. A large number of variables are simulated, including temperature, pressure, precipitation and wind. Extreme weather hazards and modes of internal variability are not explicitly coded but are emergent properties of the individual processes. Physical modelling of large waves has much in common with atmosphere–ocean modelling: one example is described by Tolman (2009), and an overview can be found in Kharif *et al.* (2009). Forecasting of tropical cyclones and of weather-related impacts is discussed at the end of the section.

NWP models are initialised with 'today's weather' using current observations of the atmosphere and ocean, then run forward for a specific time. This produces a single prediction, or realisation, of weather, which can be converted to a binary forecast for a hazard event: for example, whether the forecast period includes a storm of a given magnitude at a given location. To generate many realisations and approximate an event probability, random elements are introduced and the simulator is run several times simultaneously (Section 5.3.2.2).

Forecast lead times range from a few hours or days (short-range), to a fortnight (medium-range), to 6–12 months (seasonal). The evolution of physical quantities, such as the transfer of heat, is calculated in discrete timesteps represented with a three-dimensional grid or spherical harmonics, with spatial domain and resolution chosen according to the lead time of the forecast, timescale of the physical processes and available computing resources. Short-range forecasts are often performed for a sub-continental region, at a horizontal resolution of around 5–25 km; vertical resolution ranging from tens of metres near the surface to several hundred metres in the stratosphere; and temporal resolution of a few minutes. Seasonal forecasts are made over a global domain with climate models (Chapter 6), with a reduced resolution (around 140 km horizontally in the mid-latitudes) to reduce computational expense. Physical processes that occur on smaller spatio-temporal scales than the simulator resolution, such as convective cloud formation, are represented in parameterised form (Chapter 6). Forecasts have improved since the addition of stochastic elements in the form of randomly varying parameters or processes (Section 5.3.2.2).

Forecasting models require both initial conditions to prescribe the initial state, derived from observations and previous forecasts, and 'boundary conditions' to control the system evolution, supplied by global model simulations (Chapter 6) and 'analysis' datasets (e.g. sea surface temperatures from combining satellite and *in situ* measurements). Because the atmosphere is chaotic (Section 5.3.1.1), forecast success is heavily dependent on the

accuracy of model initialisation. NWP model simulations are therefore updated, or 'nudged', frequently by current observations in a process known as 'data assimilation' (Evensen, 2009), which takes into account the relative uncertainties of the observations and simulations. Data assimilation is an inverse method founded in Bayes theorem (Wikle and Berliner, 2007) in which NWP model predictions (samples from the prior distribution) are updated with observations to estimate the most likely atmospheric state (mode of the posterior distribution). The dimension of this state is extremely high so there are various simplifications, such as estimation of means or modes and variances rather than full probability distributions, and assumptions, such as linearity or Gaussian uncertainties, to make the problem tractable. 'Variational' methods of data assimilation (e.g. Rabier *et al.*, 2000) are gradient-based optimisation schemes, which minimise a cost function (the simplest being least squares) to find the mode of the posterior distribution; this requires an 'adjoint' model, constructed by differentiating the model equations by hand or with automatic software, which is very challenging for such complex models. 'Ensemble' methods (e.g. Evensen, 1994) use a set of about 100 simulations, each starting from slightly different initial states to approximate the mean and covariance of the prior distribution, but these are not yet used much in operational forecasting (e.g. Houtekamer *et al.*, 2005). Methods with few or no simplifying assumptions are currently only suitable for very low dimensional problems, though their efficiency is continuously being improved (van Leeuwen, 2009, 2010, 2011).

Seasonal forecasts for tropical cyclones are currently undergoing a sea change in methods and capabilities. Predictions were traditionally based on statistical analysis of past relationships with climate indicators such as Atlantic sea surface temperatures and ENSO (Section 5.3.3.3). Now a small number of forecasts use dynamical methods, applying criteria to climate model simulations for variables such as sea surface temperature, vertical shear of horizontal wind and low-level vorticity, to infer the presence of tropical storms (Vitart and Stockdale, 2001; Barnston *et al.*, 2003; Camargo and Barnston, 2009). A hybrid 'statistical-dynamical' method is described by Emanuel *et al.* (2006) in which synthetic storms are seeded randomly with tracks derived from statistical modelling of re-analysis data (Section 5.3.3) and intensities simulated by a high-resolution physical model. Dynamic modelling has only recently become comparably successful to statistical modelling at making seasonal predictions of tropical cyclone frequency (e.g. Vitart *et al.*, 2007). As computational power increases, model resolution can increase and this is likely to further improve prediction skill (Vitart *et al.*, 2007). Overviews of seasonal tropical cyclone prediction are given by Camargo *et al.* (2007) and Klotzbach *et al.* (2011).

Hazard impacts can be simulated if the important processes are understood. A well-developed hazard impact model can be chained, or 'coupled', to an NWP model to give a probabilistic assessment of loss, including real-time forecasts if the lead time is sufficiently long (days or longer), or the hazard onset slow (e.g. drought), or the impacts slow to emerge (e.g. cumulative impacts of extreme weather on vegetation: Senay and Verdin, 2003; Chapter 6). If the processes are not well understood, or the timescales short, statistical modelling of past events is used instead (Section 5.3.3.3): for example, impacts on human

health and financial loss are too challenging to describe with physically based laws, so they are forecast with statistical models or expert synthesis of observations and physical understanding (e.g. Funk, 2011).

5.3.2.2 *Uncertainty assessment in physical modelling*

Uncertainty assessment is relatively well-developed in hydrometeorogical hazard risk analysis due to a long history of weather recording and forecasting. Forecast uncertainty is a function of the observational and NWP model uncertainties. The former arise from measurement uncertainty, incomplete coverage and uncertainty in the 'observation operators', which translate measured quantities into simulated (e.g. transformation, interpolation or integration). The latter are more difficult to estimate and arise from uncertain model parameters (representing sub-grid-scale processes) and model structure (due to missing or poorly understood processes and finite resolution). Forecast uncertainty is estimated by sampling from these uncertain quantities and running different variants of the forecast simultaneously to produce an ensemble prediction.

Ensembles are generated with several types of perturbation to explore this uncertainty. Weather forecasts are critically dependent on the accuracy of the initial state, so 'initial condition ensembles' have been used in numerical weather prediction since the 1990s. Initial state perturbations are either constructed randomly, or by sampling from the observational uncertainties, or to generate the fastest growing impact on the forecast (e.g. Persson and Grazzini, 2007). Parameter and structural uncertainty are estimated by perturbing the model configuration (e.g. Houtekamer *et al*., 2009): for example, with varying resolution (Persson and Grazzini, 2007); stochastically varying parameters (Buizza *et al*., 1999; Palmer, 2001; brief overview by McFarlane, 2011); or stochastic kinetic energy backscatter (Shutts, 2005; Berner *et al*., 2009). These schemes explore both structural and parametric uncertainty simultaneously: for example, stochastic parameters allow the model to move into rarer states, such as blocking-type flow in Northern Hemisphere winters (Palmer *et al*., 2005). Another approach to estimating structural uncertainty is with a multimodel ensemble (MME), a group of models from separate research centres (e.g. Palmer *et al*., 2004; Hagedorn *et al*., 2005; Weisheimer *et al*., 2009). MMEs are mainly used for forecasts made over global domains (medium-range and longer lead times; see Chapter 6) because short-range forecasts are often made with regional models so there are fewer opportunities for comparison. Individual members of an MME are averaged together or considered as a group; they may be weighted equally or according to their degree of success in matching observations (discussed below). The mean of an MME is often found to outperform the individual members (e.g. Hagedorn *et al*., 2005; Vitart, 2006), increasingly so with ensemble size (e.g. Ferro *et al*., 2008), though it does not outperform the best member. Model structural uncertainty can also be estimated with non-ensemble approaches (Hollingsworth and Lönnberg, 1986; Parrish and Derber, 1992).

The frequency with which a given hazard appears in an ensemble forecast is treated as an estimate of its probability of occurrence: for example, a heat wave that occurs in 20% of the ensemble simulations is assigned a probability of 20%. If the NWP model is well-calibrated

(below), then over the long term such an event will occur in 20% of such forecasts. Ensemble forecasts tend to be referred to as probabilistic, but they are more accurately described as frequency distributions due to their low sample size and incomplete representation of all uncertainties; further discussion can be found in Chapter 6.

Calibration of NWP forecasts is tuning of the ensemble perturbations so that forecast frequencies match the real world over the long term. Forecasts are repeatedly tested against observations to calculate skill scores, metrics that quantify the success of the model forecast relative to a reference such as local climatology or the weather of the previous time period (e.g. Brier, 1950; Talagrand *et al.*, 1997; Persson and Grazzini, 2007). Forecast verification is an extensive field with a long history: reference texts are Jolliffe and Stephenson (2003) and the relevant chapter in Wilks (2011), and a review of recent advances is given by Casati *et al.* (2008). Over the past decades, improvements in model structure, resolution, data assimilation methods and observational datasets have led to continuous improvement in skill scores: the US National Centers for Environmental Prediction (NCEP) 'S1' metric, for example, has been recorded since the mid-1950s and shows a doubling in forecast length for a given skill score over 1–2 decades (Kalnay, 2003). Ensemble forecasts are generally underdispersed: over the long term, they are more narrowly spread than the observations. This bias may be partially corrected with post-processing methods that re-weight the ensemble members according to their performance during a training period (e.g. Gneiting, 2005; Raftery *et al.*, 2005). Post-processing methods are currently being adapted to allow spatially varying calibration (Kleiber *et al.*, 2010) and spatio-temporal correlation structure (e.g. Berrocal *et al.*, 2007), but not yet multivariate correlation.

Forecasts of extreme weather events cannot be verified, of themselves, with the usual skill scores: not only does their rarity drastically decrease the number of tests and increase the uncertainty, but the skill scores tend towards trivial limits (such as zero) with increasing event rarity (Stephenson *et al.*, 2008). Various alternatives have been proposed, either by constructing new skill scores suitable for extremes (Casati *et al.*, 2008) or by adapting extreme value theory (Section 5.3.3.1) for verification purposes (Ferro, 2007), but in general calibration is performed with normal day-to-day weather. Some extreme weather events are much harder to predict than others (Section 5.3.1.3), which means, in effect, that NWP models are less well calibrated for these events, and that the assessed probabilities must be interpreted with caution.

5.3.3 Statistical analysis

The objective of statistical analysis is to assess the probability distribution of a future hazard or its impact. Statistical methods for assessing future hazard event probabilities are only valuable to the extent that a defensible relationship can be proposed between historical and future events. For hydrometeorological events, the biggest challenge is the non-stationarity of the earth system during the instrumental period. Return periods are interpreted under an assumption of stationarity, so if the drivers that determine hazard frequency and properties

do change significantly then the predictions become less reliable. Some hazards are not significantly affected, but for most there are trade-offs if choosing subsets of the data that are approximately stationary, and difficulties if building non-stationarity into the statistical model.

Statistical analysis of hydrometeorological hazards takes many forms, according to the objectives of the analysis and the characteristics of the available data. The observational datasets analysed for extreme weather assessments are typically point data from weather stations. Other types of data include: satellite observations; re-analysis datasets such as ERA-40 (Uppala *et al*., 2005), which are continuous gridded data generated by assimilating observations into NWP model simulations; or model simulations alone, often in relation to climate change (Chapter 6). Observational datasets of hazard impacts come from a wide range of sources: in particular, government and hospital records for mortality and morbidity, and insurance losses for financial impacts. Observational data may only be available as maxima of a given time period, usually one year, known as 'block maxima' (or equivalently minima). However, block maxima are, by definition, a small subset of the observations. Other data may contain useful information in the properties of, for example, the second- and third-largest events in each time period. If all the observational data are available, it can be preferable to analyse the r-largest maxima ($r > 1$), or the exceedances over a threshold ('peaks-over-threshold', POT), or the whole time-series. The relative advantages of these are discussed in the following sections.

Three common objectives for statistical analysis are: estimating return levels, with extreme value analysis; quantifying the trend or covariates of a hazard, with regression methods; and generating 'pseudo-hazard' (also known as 'synthetic hazard') datasets, with 'weather generation' and 'catastrophe modelling'. A brief introduction to the latter two is given here.

Weather generation is the stochastic simulation of a large number of synthetic long datasets of 'pseudo-weather' (e.g. precipitation), each statistically consistent with the original observations. The main purposes of this are to analyse their characteristics (e.g. frequency of rare extremes) and to use them as inputs for impact models (e.g. crop yields), in order to supplement observational records that are short and spatially or temporally incomplete or simulations that are short and have low spatio-temporal resolution.

Catastrophe modelling is stochastic simulation of a large number of synthetic long datasets of 'pseudo-events' (e.g. hurricanes), application of these to an inventory of vulnerable quantities (usually buildings) and estimation of the damage and associated losses (usually financial; sometimes also environmental and social). Catastrophe ('cat') models were first developed in the late 1980s by three US firms (AIR Worldwide, EQECAT and Risk Management Solutions) to improve insurance risk management. After several major hurricane and earthquake events worldwide, including Hurricane Hugo in 1989 and Hurricane Andrew in 1992, these firms made their models more sophisticated and their exposure databases more comprehensive (Cummins, 2007). Other companies subsequently developed in-house models, and US federal and state government supported development of two free open-source models: HAZUS-MH (Vickery *et al*., 2000a, 2000b, 2006a, 2006b;

FEMA, 2007) and the Florida Public Hurricane Loss Model (FPHLM) (Powell *et al.*, 2005; Pinelli *et al.*, 2008). The details and assumptions of the free models are available in the peer-review literature and technical documents, because they are intended for risk mitigation, regulation and emergency preparation, but the majority of cat models are proprietary and intended for pricing insurance or reinsurance policies, so their details are not public.

Cat models are used to make probabilistic estimates of property losses: mostly residential, commercial and industrial buildings, but also forests and off-shore energy structures such as oil platforms and pipelines. These may be general assessments, such as expected losses for the following year, or specific forecasts for a recent hazard event (e.g. insured losses from Hurricane Katrina: RMS, 2005). Their comprehensive modelling of the entire hazard-to-loss causal chain – hazards, impacts, exposure, vulnerability, damage and loss – is unique in hydrometeorological hazard risk assessment. Cat models combine several modules: a hazard model (statistically based, but where possible physically motivated: Section 5.3.3.2); extensive databases that map exposure and vulnerability (such as property location, market value, reconstruction value, age, construction type and number of storeys); a damage model (structural damage as a function of hazard intensity per building type, derived from wind tunnel tests, post-event analysis and expert judgement); and a cost model (including cost of repairs/replacement, cost increases due to demand surge, business interruption, relocation and insurance policy details). Monte Carlo techniques are used to simulate tens of thousands of years of hazards, as in weather generation, which are propagated through the modules to estimate statistical properties of loss. Cat models are tailored to particular hazard types – the first were constructed for earthquakes, hurricanes and floods, but other natural hazards are now covered, including mid-latitude wind storms, winter storms, tornados and wildfires – and particular regions, such as the Gulf and Atlantic coasts of the United States. The region-specific information is not only the portfolio at risk but also local environmental conditions such as topography (for modelling precipitation), surface roughness (for wind dissipation) and location of trees (for tree debris). Cat model outputs include probability distributions of loss and exceedance probability (EP) curves for both total and insured losses. A reference text is Grossi and Kunreuther (2005).

The three objectives described above – return levels, covariates and 'pseudo-hazards' – are discussed here according to the characteristics of the data: stationary and independent (Section 5.3.3.1), correlated (Section 5.3.3.2) and non-stationary (Section 5.3.3.3). A reference text for extreme value modelling is Coles (2001). An overview of techniques for correlated or non-stationary hydrometeorological observations is given by Khaliq *et al.* (2006).

5.3.3.1 Stationary independent data

If observations can be treated as independent and stationary, estimation of the frequency and intensity of extremes is greatly simplified. How often this may be possible is discussed at the end of the section.

Under these assumptions, and if the record is sufficiently long, the extremes of continuous data may be characterised with the extreme value limit theorem, which is analogous to the
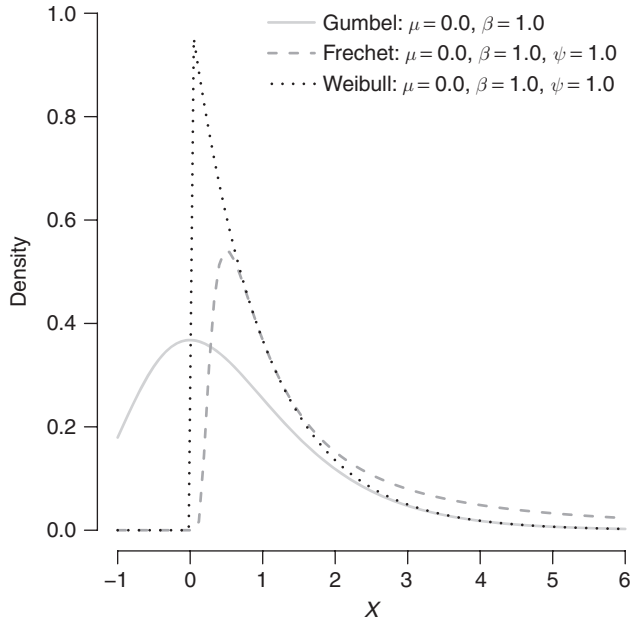
Figure 5.4 Examples of extreme value distributions, with values given for the location (μ), scale (β), and shape (ψ) parameters.

central limit theorem. This states that the renormalised (explained below) block maxima of random observations from the same distribution tend asymptotically, if they converge at all, to one of the generalised extreme value (GEV) family of distributions: Gumbel (exponential-tailed), Fréchet (heavy-tailed) or negative Weibull (bounded, light-tailed). Examples of these are shown in Figure 5.4. The *r*-largest maxima in a block can equivalently be described with a more general form of the GEV (see Coles, 2001). The extreme value limit theorem also states that the amounts by which a threshold is exceeded tend asymptotically to one of the generalised Pareto distribution (GPD) family of power law distributions. In the case of *r*-largest and POT, there is a bias–variance trade-off in which larger (typically longer) samples produce tighter predictions, but may introduce bias through more serious violations of modelling assumptions. The choice may be aided with a graphical method (mean residual life plot: Coles, 2001) or sensitivity studies (see below).

The strength of extreme value theory (EVT) is that it provides tools to estimate return levels without making a-priori assumptions about the probability distribution from which the observational data or their extremes are sampled. The limit theorem removes the need to model the full dataset, and the distribution of the extremes can be inferred empirically by estimating the characteristic parameters from the data. The GEV distribution has parameters for location, scale and shape, which control the mean, asymmetry and rate of tail decay of the distribution, respectively: a positive shape parameter corresponds to the Fréchet, zero to Gumbel and negative to the negative Weibull. Early extreme value analyses did make

a-priori assumptions about which of these applied, constraining the range of the shape parameter, but this is generally thought unnecessary and undesirable unless there is a strong physical motivation. Renormalisation is simply subtraction of the location parameter and division by the scale (Coles, 2001). The GPD has scale and shape parameters: shape equal to zero corresponds to the exponential distribution.

The GEV and GPD are the dominant model choices, but other distributions are occasionally used, such as the generalised logistic (GLO) distribution (e.g. Kyselý and Picek, 2007) or the more general kappa family of distributions (e.g. Park and Jung, 2002), of which the GEV, GPD and GLO are members. Wave heights above a threshold may be fitted with a modified Rayleigh distribution, one of the Weibull family (Tucker and Pitt, 2001). EVT does not apply to discrete distributions: count data, such as number of tropical cyclones or threshold exceedances, are usually modelled with a Poisson distribution or similar (e.g. Jagger and Elsner, 2006). A parametric model (distribution or family of distributions) is almost always specified, because it enables extrapolation to rarer extremes than are present in the data, but non-parametric approaches are available (e.g. Beirlant *et al.*, 2004).

Most approaches are frequentist, treating the parameters as fixed but unknown. One of the most commonly used methods of inference is maximum likelihood (ML) estimation, a broad range of techniques including the method of least squares for finding single 'best' parameter values with confidence intervals. ML estimation is very often preferred because it is intuitive, many software tools are available and it is flexible: for example, it is easy to include covariates (Section 5.3.3.3). For small samples, ML may give a poor estimate (Engeland *et al.*, 2004) so moment-based approaches such as probability-weighted moments or L-moments (Storch and Zwiers, 2003) might be more appropriate, though moments do not always exist for extreme distributions. Model fitting may be repeated for sensitivity studies: for example, to assess the stability of parameter estimates to the choice of threshold or r value. Once the parameters are estimated, model 'goodness-of-fit' is assessed with graphical or other tests. If the model is thought to describe the data sufficiently well, it is used to estimate return levels, usually for more extreme quantiles than are present in the observations. Uncertainty in the return level estimates is assessed using likelihood or bootstrapping methods. Parameter and model uncertainty assessment are discussed in Section 5.3.3.4.

A judgement must be made about whether the data may be considered stationary. Records of extremes that vary seasonally may be treated as stationary by using the annual maxima, but long-term, interannual non-stationarity might seem hard to avoid, given the earth's climate is never at equilibrium. However, stationarity might be approximated by taking logarithms of the ratio of successive observations (Coles, 2001), or assumed if the observed quantities are not closely linked with global atmospheric temperature (e.g. Wimmer *et al.*, 2006), or the record relatively short (e.g. Baccini *et al.*, 2008), or the spatial domain small (e.g. Kyselý and Picek, 2007). The trade-off in using a spatial or temporal subset to approximate stationarity is that it leads to larger statistical parameter uncertainties and thus less reliable probability assessments, especially for large-magnitude hazards. Another important point to consider is whether changes in the measurement process – for example in

spatial coverage, fraction of events recorded or systematic error – have created a false impression of non-stationarity (e.g. surface temperatures: Brohan *et al*., 2006; rain gauges: McGregor and MacDougall, 2009; hurricanes: Vecchi and Knutson, 2011).

Non-stationarity is present to different degrees in different hazard observational records. The number of warm days (globally) and heat waves (in many regions) are thought to have increased since 1950, as have the number of heavy precipitation events and drought intensity and length in some regions (IPCC, 2011). Storm intensity in Europe does not seem to show a significant trend during the twentieth century (Della-Marta *et al*., 2009). Hurricane frequency is thought not to have increased globally during the twentieth century, but there is disagreement about whether their intensity has increased with global mean temperature or is varying on natural cycles (Emanuel, 2005; Webster *et al*., 2005; Pielke, 2005, 2006; Anthes *et al*., 2006; Landsea *et al*., 2006; Landsea, 2007; IPCC, 2011). Non-stationarity of the hurricane measurement process is a contributing factor in this debate.

Economic losses from extreme weather events have increased, largely due to increasing exposure and assets (IPCC, 2011). Losses due to hurricanes have been concluded to be stationary after normalising for inflation and increasing population, wealth and (in the case of the United States) settlement of coastal areas for US losses from 1900–2006 (Pielke *et al*., 2008) and global losses from 1950–2005 (Miller *et al*., 2008), though there may be a trend in the United States since 1970 due to increased hurricane activity in this region (Miller *et al*., 2008; Schmidt *et al*., 2009). It is difficult to account for all covariates, such as loss reporting and vulnerability, in these studies (Miller *et al*., 2008).

A general note of caution on the assumption of stationarity: a failure to reject the null hypothesis that the historical process is stationary may simply reflect the low power of the test statistic for identifying a non-stationary process that may behave differently in the future. If there are underlying physical reasons for believing that the data are stationary, this may be more robust.

5.3.3.2 Correlated data

Correlations in data may occur in the form of multivariate, spatial or temporal dependencies. These are common in hydrometeorological hazards due to the teleconnections and feed-backs that link places and aspects of climate with each other, the large spatial extent of atmospheric and ocean currents and the temporal persistence of meteorological and environmental phenomena.

There are analogues of EVT for multivariate extremes (such as heat waves, droughts and storms), including component-wise block maxima and threshold excesses (Coles, 2001). However, this involves sequential ordering of each dimension: results depend on the ordering and there may not be an obvious choice for a given problem. An alternative is the copula approach, in which the joint distribution for continuous random variables is broken down into the individual, 'marginal' distributions and a copula function that models their dependence structure. A family of copulas is chosen (e.g. Gaussian) and their parameters estimated along with those of the marginal distributions. A general reference for copulas is given by Nelsen (2006), their application to risk management by McNeil *et al*.

(2005) and to natural hazards by Salvadori *et al*. (2007). Schölzel and Friederichs (2008) give a useful overview and call for copulas to be more widely used in weather and climate research.

Hydrometeorological observations are very often correlated in space and time because of spatially extensive or temporally persistent weather systems and modes of natural variability. In the case of continuous variables, this can lead to a very heavy-tailed distribution for which the GEV and GPD are no longer valid. Block maxima might be treated as independent, if the dependence in the full dataset is not too long-range, but threshold exceedances cannot due to the 'clustering' of events that exceed a given threshold. This may be dealt with by processing the data to approximate temporal independence or by adjusting the assumed frequency model to account for the dependence.

The simplest approach for approximating temporal independence is to apply a simple filter. Brown *et al*. (2008), for example, thin their temperature observations by using only the maximum threshold exceedance in each consecutive ten-day window. This fixed-window approach fails if there are some particularly persistent clusters in the data. More often used is a tailored pre-processing, known as 'declustering', which identifies individual clusters in the data in order to model their maxima. The most common of these is 'runs declustering', which uses interexceedance time (a minimum gap length between clusters; e.g. Della-Marta *et al*., 2009). These methods necessitate a bias-variance trade-off between the number of observations and their independence, an obvious disadvantage for datasets of rare events. Their design may also be rather subjective. An alternative is decorrelation, or 'pre-whitening', in which dependence is modelled and removed. In general this is only appropriate if there are physical motivations for a particular correlation structure, and care must be taken to avoid removing an underlying trend (Storch, 1995; Khaliq *et al*., 2006).

A simple approach to incorporating dependence in the model is to inflate the variance: for example, by replacing a simple Poisson model of count data with one that includes a dispersion parameter (e.g. negative binomial, or the generalised linear model family), thereby allowing for overdispersion (e.g. Vitolo *et al*., 2009). Another commonly used method is to empirically estimate the 'extremal index' and use this to modify the parameters in the GPD (Coles, 2001). The extremal index is the reciprocal of the 'limiting mean cluster size': for example, an extremal index of 0.2 indicates a mean cluster size of five exceedances, where 'limiting' refers to clusters of exceedances of increasingly high thresholds. Extremal index is sensitive to cluster definition, but fortunately this does not always propagate to the estimated return levels (Coles, 2001). An alternative is to model the correlation explicitly and incorporate this in the statistical inference: for example, Baccini *et al*. (2008) use generalised estimating equations (GEE) for count data to allow the inclusion of correlation structure. Wavelet and kernel density estimation methods have also been proposed for dependent data (Khaliq *et al*., 2006).

Weather generation has historically focused on precipitation, for which temporal persistence is important. In parametric weather generation, time dependence is described with an autoregressive or a clustering model (e.g. Jones *et al*., 2009; Kyselý, 2009; Nicolosi *et al*., 2009). Non-parametric methods, which sample from historical observations, must preserve

temporal structure: for example, by taking multi-day sections of data. Models with spatial correlation have been developed (Wilks and Wilby, 1999; Mehrotra *et al.*, 2006), but in general weather generation is location-specific (e.g. Jones *et al.*, 2009).

In contrast, catastrophe modelling of hurricanes attempts to quantify their spatial extent, direction of movement, evolution through time and spatio-temporal propagation of their impacts. Spatio-temporal evolution is important for estimating damage: a low-category hurricane that lingers can be more destructive than a high-category one that moves swiftly though. This evolution can be difficult to infer from individual wind speed observations, if they are even available (anenometers often fail in high winds), but is straightforward from individual hurricane tracks. Cat modelling thus has a different philosophical grounding to EVT modelling: it attempts to capture the space-time behaviour of entire event systems (hurricane genesis, landfall and decay) and derive wind speeds from these, rather than characterising the wind speed point data directly.

Cat modelling of hurricanes, though physically motivated where possible, is fundamentally statistical analysis of historical data. Hall and Jewson (2007, 2008) describe various methods for stochastically generating synthetic storm data by calibrating with, or sampling from, historical observations of their location, intensity, direction and speed. Separate statistical models are used for each aspect of the storm: for example, a spatial Poisson process for genesis; a random walk for deviation from mean tracks; and, more recently, regression against time of year and sea surface temperature for evolution of intensity. The impact of the hurricane is then propagated with a physically motivated statistical model to generate region-wide wind fields (with wind magnitudes derived from hurricane direction, circular motion and speed of decay) and downscaled (Chapter 6) to derive winds at point locations; other quantities such as precipitation are also modelled. Correlation is incorporated by the sampling or modelling of entire hurricane systems through time and with model parameters that vary smoothly in space. Cat models typically assume individual hurricane events are temporally independent, although research methods are being developed to account for clustering. They also assume stationarity, but tend to be used for very short time horizons or else observe that the loss uncertainty dominates the hazard uncertainty.

5.3.3.3 Non-stationary data

Most datasets of hydrometeorological extremes are non-stationary. If a near-stationary subset cannot be selected, one solution is to build non-stationarity into the statistical model, but this raises the issue of how to ensure the model holds in the past and continues to hold in the future under changing forcings. Another solution is to regress the hazard or impacts onto covariates, but this also requires assumptions about the stability of these relationships into the future. Statistical modelling approaches to non-stationarity are discussed in this section; physical modelling approaches are discussed in Chapter 6.

Some kinds of non-stationarity are more difficult to deal with than others. The seasonal cycle is an important time-varying influence on many hazards, but a relatively predictable one. The effect of a mode of natural variability on a hazard can be quantified, and if it is long-range (in time) can be used as a predictor variable for forecasting. However, the long-term

trends of climate change, and other driving factors such as land use and societal change, are only somewhat predictable in the near-term and much more uncertain after a few decades. For the hazards and impacts that are influenced by changing drivers, neither an assumption of stationarity nor a model of recent non-stationarity are likely to be valid for long.

There are a variety of strategies available to incorporate non-stationarity in a statistical model of historical observations. Typically these focus on changes in location (the mean of the data, or the location parameter of the distribution). They may also include the scale or other parameters, though it is more difficult to estimate the shape parameter when it is allowed to vary (Coles, 2001). Selection of the extremes may be affected: for example, in POT analysis the threshold of exceedance might be varied with time (e.g. Brown *et al.*, 2008; Della-Marta *et al.*, 2009). The simplest approach is to divide the data into subsets that can be assumed stationary and estimate the model parameters separately for each. The subsets may be determined by non-overlapping time windows (e.g. Coles *et al.*, 2003), moving time windows (e.g. Zhang *et al.*, 2001), an external factor such as 'high' or 'low' values of an index of climate variability (e.g. Jagger and Elsner, 2006) or abrupt jumps ('change points') identified with regression, non-parametric (e.g. Li *et al.*, 2005) or Bayesian methods (e.g. Elsner *et al.*, 2004; Zhao and Chu, 2010). An overview of change point detection is given by Kundzewicz and Robson (2004). Dividing the data into subsets is also used for assessing changes in return period and in generating pseudo-weather datasets for the present day and long-term future (Chapter 6). A more flexible approach to discrete subsets is to model the dependence of the parameters on time: as a function of the year, the seasonal cycle or a combination of these (e.g. Ribereau *et al.*, 2008; Kyselý, 2009; Menéndez *et al.*, 2009).

Non-stationarity may be diagnosed using non-parametric tests based on ranking or resampling of the data such as the Kendall's tau (e.g. Kunkel *et al.*, 1999; Zhang *et al.*, 2001; Chavez-Demoulin and Davison, 2005). This avoids the need to propose a model for the data, but does not allow extrapolation outside the data nor identify the shape of the trend. Kundzewicz and Robson (2004) outline some commonly used non-parametric tests of trends in hydrological data. A proposed semi-parametric approach is 'local likelihood', in which parameters are estimated separately at each time using the whole dataset and re-weighted so that nearby points in time assume more importance (e.g. Ramesh and Davison, 2002); this may be helpful in choosing a suitable parametric form (Khaliq *et al.*, 2006).

It may be desirable to remove non-stationarity to de-trend the data. Eastoe and Tawn (2009) point out that there can be problems in fitting data if the exceedance threshold is allowed to vary with time, so they propose a pre-processing approach where the trend is modelled and removed. Pre-processing, also known as 'pre-whitening' (as is decorrelation, Section 5.3.3.2), is discussed by Chatfield (2004).

A hazard, or parameters of a distribution such as the GEV, may be modelled as a function of other variables. These regression approaches typically aim to assess the influence of climatic factors on hazards or the influence of hazards on vulnerable entities. Covariates may be trends (such as global mean temperature) or oscillations (such as modes of natural variability), though over short timescales it may be difficult to distinguish between these.

The parameters of the regression model are estimated with ML or other methods (Sections 5.3.3.1, 5.3.3.4).

Regression of hazards on preceding climatic conditions can be used for near-term forecasting. This has been particularly important for tropical cyclones on seasonal timescales, for which physical modelling is challenging due to the small spatio-temporal scales of cyclogenesis and the initial condition uncertainties. Counts of hurricanes, for example, have been modelled using Poisson regression or the more flexible GLM with respect to sea surface temperatures (AIR Worldwide Corporation, 2006) or metrics such as the ENSO, NAO, SOI (Southern Oscillation Index), Atlantic Multidecadal Oscillation, and Sahel Rainfall Index (Elsner, 2003; Elsner and Bossak, 2004; Elsner and Jagger, 2006). Hurricane wind speeds have been modelled with respect to global mean temperature (Jagger and Elsner, 2006). Camargo *et al.* (2007) outline predictor variables for statistical seasonal hurricane forecasts from different meteorological agencies around the world. Forecasts of hurricane frequency are increasingly made with dynamical models or statistical-dynamical methods, but forecasts of their intensity and landfall are currently more successful from statistical methods. Regression onto modes of internal variability has also been used for mid-latitude storms (e.g. Mailier *et al.*, 2006; Vitolo *et al.*, 2009) and heat waves (e.g. Rusticucci and Vargas, 2002).

Regression-type approaches are important in impacts modelling, where the physical mechanisms are often poorly known or difficult to predict. It is used, for example, to quantify the relationship between extreme temperature or precipitation and human mortality (Githeko and Ndegwa, 2001; Sheridan and Kalkstein, 2004; Baccini *et al.*, 2008), and between hurricanes and economic losses (Saunders and Lea, 2005).

Regression onto modes of natural variability is a space-for-time substitution, deducing a relationship from recent spatial patterns and using it for the future. It can also be viewed as a device for handling spatial dependence. That is, if a hazard can be seen to depend on a large-scale feature such as the NAO, spatial structure can be modelled by quantifying the dependence on the NAO in each location, treating each location as conditionally independent, then inducing spatial dependence through variation in the NAO. For assessments of the far future, regression approaches do not avoid the problem of non-stationarity. They rely on an assumption of the stability of these relationships (e.g. teleconnections) even though they may be significantly altered by changing boundary conditions.

Non-stationary spatio-temporal modelling is very challenging, especially for extremes. This is a relatively new field in statistics, and some of the simpler approaches are rather ad hoc, so it is important to be cautious in interpreting the results. Ribereau *et al.* (2008: 1037), for example, warn that 'Extrapolating the trend beyond the range of observations is always a delicate and sometimes dangerous operation (since we assume that the trend will remain the same in the future), especially when dealing with extremes. Hence high return levels in a non-stationary context must be interpreted with extreme care.'

5.3.3.4 Uncertainty assessment in statistical analysis

Statistical analysis uncertainty can be divided into those from inputs (measurements), model parameters and model structure. Measurement uncertainties vary in space and time due to

the changing coverage and quality of observation networks (e.g. surface temperatures: Brohan *et al*., 2006; rain gauges: McGregor and MacDougall, 2009; and hurricanes: Vecchi and Knutson, 2011). Iman *et al*. (2005a, 2005b) describe the effects of perturbing the inputs of a hurricane catastrophe model.

Parameter confidence intervals are commonly estimated with likelihood-based methods (such as the 'delta' method, which relies on asymptotic and regularity assumptions, or 'profile likelihood', which is generally more accurate but requires more computation; Coles, 2001) or bootstrapping methods (either non-parametric, by resampling from the original data, or parametric, by resampling from the fitted distribution; Davison and Hinkley, 1997). Alternatively, parameter probability density functions can be estimated with Bayesian methods, which are becoming more popular in extremes analysis because they allow prior information to be incorporated, estimate the entire distribution (or summary characteristics such as moments or quantiles), do not require asymptotic and regularity assumptions and are flexible and transparent, This enables assimilation of disparate information (e.g. Jagger and Elsner, 2006; Baccini *et al*., 2008), and makes it easier to 'keep track' of all sources of uncertainty (e.g. Coles *et al*., 2003). Pang *et al*. (2001) give an introduction to two of the simplest computational methods in the context of extreme wind speed analysis. Generalised maximum likelihood (El Adlouni *et al*., 2007) has been suggested as a compromise, reducing computational expense by focusing on the mode of the posterior distribution.

Quantification of uncertainties can be sensitive to the chosen subset of data and method: for example, confidence intervals may be smaller for *r*-largest maxima than block maxima (due to the larger dataset), but care must be taken that the model is still valid with the lower threshold; confidence interval estimates for some distributions may be very dependent on the method (e.g. short-tailed distribution for storms: Della-Marta *et al*., 2009); non-parametric bootstrapping may underestimate uncertainties relative to parametric boot-strapping (Kyselý, 2008); and likelihood-based methods may give unphysical parameter estimates (Pang *et al*., 2001).

Structural uncertainty in statistical analysis is associated with the choice of distribution or model and whether the necessary assumptions for that model are valid (e.g. independence, stationarity and regularity). Models are usually used to extrapolate to rarer extremes than are in the observations; GEV-based estimates of uncertainty for extrapolated values should be considered lower bounds (Coles, 2001). Robust testing of model reliability is essential. Goodness-of-fit, the success of a model in describing observations, is assessed using graphical methods such as quantile plots (Coles, 2001; Beirlant *et al*., 2004) or statistical tests (Coles, 2001); the properties of these tests vary, which can affect the results (e.g. Lemeshko *et al*., 2009, 2010).

In extreme value analysis, data selection affects model structural uncertainty: *r*-largest and peaks-over-threshold analyses use more of the observations than block maxima, and therefore provide more information with which to check the model, but if *r* is too large or the threshold is too low the model assumptions may no longer be valid. Long record length is required for the asymptotic limit, but a short record may be required for stationarity.

5.4 Risk management

Risk management comprises the steps taken to assess, mitigate and cope with risk in the periods between hazard events (monitoring and planning: Section 5.4.1) and during and after hazard events (communication, response, recovery and analysis: Section 5.4.2). Aspects of risk management that are relevant to climate change adaptation are discussed in Chapter 6.

5.4.1 Hazard monitoring and planning

Monitoring of hazards is the real-time forecasting and early warning systems of hazard events or their impacts (Section 5.4.1.1). Planning for hazards encompasses developing tools to enable rapid decision-making, such as event triggers or emergency procedures (Section 5.4.1.2), and making long-term decisions, such as land-use planning, regulation of the private sector or hedging financial risk (Section 5.4.1.3).

5.4.1.1 Monitoring and warning systems

The purpose of monitoring systems is primarily to inform decision-making before and during a hazard event. They provide information that can be translated into event probabilities, and are often assimilated into NWP models that generate warnings and trigger actions (such as communication or emergency response) when the forecast probability for a hazard of a given magnitude exceeds a threshold. Monitoring networks are well developed for most hydrometeorological hazards due to longstanding meteorological services around the world. Individual agencies such as the UK Met Office (UKMO) issue daily weather forecasts that include warnings of extreme precipitation, storms, heat waves and severe marine conditions. They also contribute to regional warning services such as the European Meteoalarm (http://www.meteoalarm.eu) and offer consulting services for users that require non-standard forecasts.

Monitoring systems are continuously updated with improved understanding of the underlying processes (e.g. for droughts, tropical cyclones) and links with impacts (e.g. for heat waves). Heat wave monitoring, for example, has traditionally been based on daily maximum temperatures, but since heat wave disasters such as the 1990s Chicago and 2003 European events a number of cities have introduced warning systems that take account of other conditions linked with excess mortality and ill health. The Heat Index of the NOAA NWS includes both temperature and humidity, and several warning systems use spatial synoptic classification (Sections 5.2.1.2, 5.4.1.2).

Lead times govern the effectiveness of early warning systems in minimising hazard impacts. Even if the inception of some hazards cannot be accurately simulated, the conditions under which they might form are known and this can be used to set a threat level. Tornado lead times are very short, up to 15–20 minutes, and warning systems are only operational in a few countries, but the atmospheric conditions in which a tornado is formed can be forecast several hours in advance, which enables people to move nearer to shelters

(United Nations, 2006). The number of tornado-related deaths significantly dropped in the United States during the last century, mainly as a result of the Doppler Radar Network (United Nations, 2006). Tropical cyclones such as hurricanes are monitored globally by the WMO Global Tropical Cyclone Warning System, with lead times ranging from 24 hours to several days (United Nations, 2006). A long lead time is necessary for large-scale evacuations to avoid widespread loss of life, but not sufficient: evacuation orders were given two days before Hurricane Katrina made landfall in Louisiana, but they were voluntary. This, along with other failings, led to avoidable deaths (Section 5.4.2.4).

For longer timescales, predictions of the overall severity of the coming tropical storm season are made with methods from across the spectrum of approaches: statistical, historically calibrated; the new generation of high-resolution dynamical models; and hybrid statistical-dynamical methods (Sections 5.3.2, 5.3.3; Camargo *et al.*, 2007). Severity is measured in terms of, for example, frequency and ACE, and expressed as summaries (e.g. median, spread and long-term mean) or probabilistic forecasts with uncertainties (Section 5.4.2.1).

Drought is a multivariate and slow-developing hazard, so warning systems are more complex and less well-developed. Droughts do not require rapid response, but the continuation of an existing drought can be assessed with an NWP model in seasonal prediction mode. Droughts are typically monitored by national or regional centres in association with meteorological agencies, using observations of several quantities including precipitation, streamflow, groundwater levels and snow pack, together with historical observations and climate model simulations (United Nations, 2006). These are combined into multi-index (Section 5.2.1.3) drought monitors that include quantitative and graphical tools. The US National Drought Mitigation Center (NDMC), for example, together with the NOAA and the US Department of Agriculture, produces the US Drought Monitor based on the PDSI, SPI and other metrics. A Global Drought Monitor based on PDSI, SPI and the NDMC drought categories is produced by the University College London (UCL) Hazard Research Centre.

Impacts of hazards may also be monitored to aid decision-making and, in the longer term, policy-making. During summer months, for example, the UK Health Protection Agency monitors heat wave health impacts using the number of calls to the National Health Service advice phoneline and the number of consultations with family doctors ('Heatwave plan for England 2009', available from http://www.dh.gov.uk). In the United States, the NDMC monitors drought impacts with the Drought Impact Reporter, which shows the number of reported impact events by state and type (agriculture, water and energy, environment, fire and social). For regions vulnerable to famine, the Famine Early Warning Systems Network (FEWS NET) monitors a number of driving factors such as rainfall and vegetation to provide early warning for issues of food security.

Although monitoring and warning systems for hydrometeorological hazards are relatively well developed, a United Nations report commented that collaboration between the hydrometeorological community and the disaster risk management and response communities could be improved (United Nations, 2006).

5.4.1.2 Rapid decision-making

Quantitative thresholds have been defined to aid decision-making during an event for many of the hydrometeorological hazards. Heat waves, in particular, have clearly defined procedures. Some are based only on temperature, while others include factors important to health identified in statistical analysis of past observations. In the United States, for example, alert procedures are triggered when the NWS Heat Index (Section 5.2.1.2) is expected to exceed 105–110°F (41–43°C), depending on local climate, for at least two consecutive days. The UK 'Heatwave plan for England 2009' has four levels of decision-making and action, each a list of responsibilities for government departments and authorities. The first is triggered by date (1 June), outlining general preparedness and urban planning measures; the second is triggered by a UKMO forecast of 60% probability of temperatures exceeding regional impact thresholds for maximum and minimum temperatures on at least two consecutive days; the third when those thresholds are exceeded; and the fourth when the heat wave is severe and prolonged. Regional maximum and minimum temperature thresholds range from 28°C and 15°C for the north-east to 32°C and 18°C for London. This system draws on the World Health Organization's 'EuroHEAT' project (Michelozzi *et al*., 2007; Baccini *et al*., 2008).

Triggers based on synoptic heat wave classification (Section 5.2.1.2) and forecast impacts are used by the Toronto Heat Health Alert System. If an 'oppressive air mass' is forecast with a probability of excess mortality exceeding 65%, various actions take place, including: distribution of information through media, government and non-profit organisations; direct contacting of vulnerable individuals; opening of a Heat Information Line; and daytime opening of homeless shelters. If the probability of excess mortality is predicted to be greater than 90%, additional actions occur such as the opening of community 'Cooling Centres' and the extension of swimming pool opening hours.

Risk management for droughts also include quantitative thresholds. In the UK, the Environment Agency produces drought plans for water companies outlining a range of trigger levels based on drought status and calendar date, with corresponding actions to manage supplies and demand (UK Environment Agency, 2009).

Severe events trigger emergency procedures, which may include mandatory orders to protect public health or services. One example is the Assistance to Shelter Act introduced in 2009 by the Legislative Assembly of British Columbia, which states that a police officer may transport a person they consider to be at risk to an emergency shelter 'using reasonable force if necessary'. In the case of severe drought in the UK, interruptions to domestic water supply can be authorised by emergency drought order (EDO); these have been used three times in England and Wales since 1945, most recently in 1976 (UK Government Cabinet Office, 2008). For some catastrophic hazards such as hurricanes, local and national governments have the authority to order mandatory evacuation, though in practice this is not, or cannot be, enforced (Fairchild *et al*., 2006; Townsend, 2006). Mandatory orders are not only given by the state or police; in the UK, restrictions on non-essential water use, such as hosepipe and sprinkler bans, may be imposed during droughts by the private water

companies without permission from the government or Environment Agency – these may be sufficient to avoid severe impacts (UK Government Cabinet Office, 2008).

Training exercises are used to inform and develop emergency plans. The US Federal Emergency Management Agency (FEMA) funded and participated in a five-day hurricane disaster simulation exercise in the New Orleans area called 'Hurricane Pam' about a year before the real Hurricane Katrina struck the same location. The hurricane category was the same and the simulated impacts very similar, including large-scale evacuation, breached levees and extensive destruction (US House of Representatives, 2006). Hurricane Pam was not a total success in preparing officials for Hurricane Katrina (Section 5.4.2.4), but the medical contingency plan it yielded, though unfinished, was judged 'invaluable' to the response effort (US House of Representatives, 2006).

5.4.1.3 Long-term planning

Long-term planning for risk management should ideally aim to use 'low-regrets' options (IPCC, 2011) that address a wide range of extreme events and other societal benefits such as improved livelihoods, sanitation, infrastructure and sustainability. Some examples are: for heat waves, improved social care networks for vulnerable groups; for droughts, improved irrigation efficiency and use of drought-resistant crop varieties; and for hurricanes, adoption and enforcement of building codes (IPCC, 2011).

Planning regulations exist to enforce good risk management in the private and public sectors and minimise losses for vulnerable entities and infrastructure. An important area of long-term planning is the setting of design codes. Buildings and other structures must be constructed to withstand extreme wind, and extreme waves and currents if offshore. Regional variations in building code levels and their enforcement can significantly affect vulnerability (RMS, 2005). Design codes are almost always expressed in terms of return values and thus inherently assume stationarity of risks. For safety reasons, the intended design life of a structure is invariably much shorter than the design value. The most commonly used design return values are 50 or 100 years. Ship certification agencies such as Lloyds and Norske Veritas demand that vessels can withstand 50-year return wave conditions, and limit their certification to those sea areas and seasons where this can be guaranteed.

Insurers manage risk using the predictions of catastrophe models (Section 5.3.3). A simple example given by RMS (2008) is summarised here. If a cat model (or the weighted or unweighted mean of several models) estimates the 99.6th percentile probability of exceedance at US$40 million insured losses for a given year, but the insurance company only has capital to cover US$30 million losses, it must reduce the portfolio at risk. Some options for this are: increasing premiums, refusing coverage or hedging the risk with reinsurance or catastrophe bonds (insurance-linked securities related to natural catastrophes). 'Cat' bonds are increasingly seen as more desirable than reinsurance because they provide access to a very large pool of capital and most investment portfolios have little correlation with natural catastrophes (RMS, 2008). More details on risk management with catastrophe models are provided by Grossi and Kunreuther (2005).

Insurance industry risks are regulated by external public bodies to help protect the public and public funding from profiteering or collapse. The Florida Commission, for example, certifies catastrophe models to ensure they give plausible predictions and that insurers have sufficient capital for the predicted exceedance probabilities. The 'Solvency II' regulation treaty is a recent development in the European insurance industry with global repercussions. European insurance and reinsurance companies must demonstrate to European Union regulators their solvency under the risk they retain, and catastrophic risk from hydrometeorological hazards is a major contributor. 'Solvency II' is driven by the European Union, but companies with global portfolios must estimate their global risk. There are similar initiatives in the Japanese and American insurance and reinsurance markets. A further example of insurance industry regulation is described in Section 5.4.2.4.

5.4.2 Hazard communication, response and recovery

5.4.2.1 Communication

Risk communication for extreme weather hazards might have some advantage over other hazards because the public are exposed to daily weather forecasts, though familiarity could breed overconfidence in interpretation. There is evidence that the public infer their own probabilities from deterministic forecasts based on their past experiences, such as a decrease in forecast skill with increasing lead time, or a reduced skill for rainfall relative to temperature (Morss *et al.*, 2008; Joslyn and Savelli, 2010), but this blurs the distinctions between forecaster, decision-maker and user. Alternatively, a probabilistic forecast can be provided: since the 1960s, the NWS (known at the time as the Weather Bureau) has supplied probability of precipitation (PoP) forecasts. There is evidence to suggest that users make better decisions if forecast uncertainties are given (Roulston *et al.*, 2006; Roulston and Kaplan, 2009).

Extreme weather warnings summarise ensemble forecasts (Section 5.3.2.2) with varying degrees of information content (see Stephens *et al.*, 2012). The 'probability' or 'risk' of a hazard essentially corresponds to the fraction of the ensemble that predicts a given event, though additional processing and interpretation by human forecasters is usually involved. The fraction may be expressed in categories: the UKMO, for example, issues 'Severe Weather Warnings' using a traffic-light system with colours ranging from green for less than 20% ('Very low') risk to red for greater than or equal to 60% ('High') risk. A similar approach is a spatial map of probability contours, such as the NOAA Storm Prediction Center forecasts for wind storms, tornados and hail (Figure 5.5); in the wind storm maps, hatched areas correspond to 10% or greater probability of wind gusts 65 knots or greater within 25 miles of a point. The NOAA provides a 'conversion table' alongside these probability forecasts, with which users can convert the numbers to risk categories 'slight', 'medium' and 'high'.

For continuous variables (rather than discrete events such as storms), full ensemble forecasts can be displayed as 'spaghetti diagrams'. An example of this is provided by the NOAA NWS, though on the NCEP Environmental Modeling Center website rather than the main NWS forecast site. The ensemble members are shown as individual contours, which
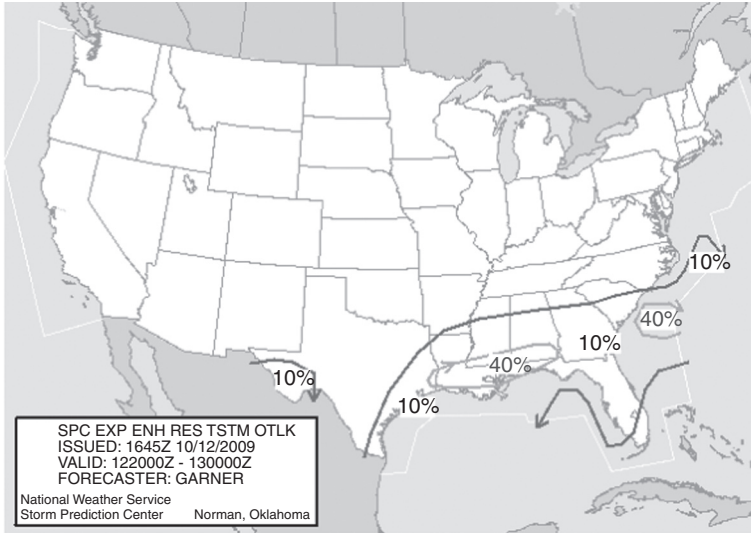
Figure 5.5 An example of the NOAA Storm Prediction Center Experimental Enhanced Resolution Thunderstorm Outlook: each forecast contains 10%, 40% and 70% contours for the probability of thunderstorms during the forecast period. From the National Weather Service.

diverge through time (Figure 5.6). Spaghetti diagrams are analogous to the 'smiley face' representations of risk often used in healthcare, because they show 'multiple possible futures' (Edwards *et al*., 2002).

However, these forecasts of ensemble frequencies generally do not contain estimates of uncertainty or model skill. Tropical storm forecasts issued by the NOAA are one exception: their uncertainty is communicated in a variety of graphical representations, of which the most well-known is the National Hurricane Center (NHC) 'cone of uncertainty' for storm tracks. This includes: present storm location and type, forecast track line, a white-filled cone representing average forecast error out to three days, a black cone outline representing forecast error for days four and five, and coastlines under a watch/warning (Figure 5.7). The cone is constructed from a series of circles, each of which represents a forecast period (e.g. 36 hours), with the size of each circle set to enclose 67% of the previous five years' official forecast errors. The cone of uncertainty is intended for real-time decision-making by the public, government agencies and the private sector, and is widely used and largely success-ful. However, many users do misinterpret it. Some interpret the central track line as the only area of impact, while others interpret the cone as the area of impact rather than of storm track uncertainty (for more detail, see Broad *et al*., 2007; Stephens *et al*., 2012). This has serious implications because (by design) the cone only predicts the central path correctly about two-thirds of the time. Broad *et al*. (2007) show that 'more' is not always better than 'less' in the case of forecasting, and that uncertainty in one quantity (track location) may be misinter-preted as certainty in another (impact region). New visualisation options have been suggested (Figure 5.7), but the cone of uncertainty has not, as yet, substantively changed.
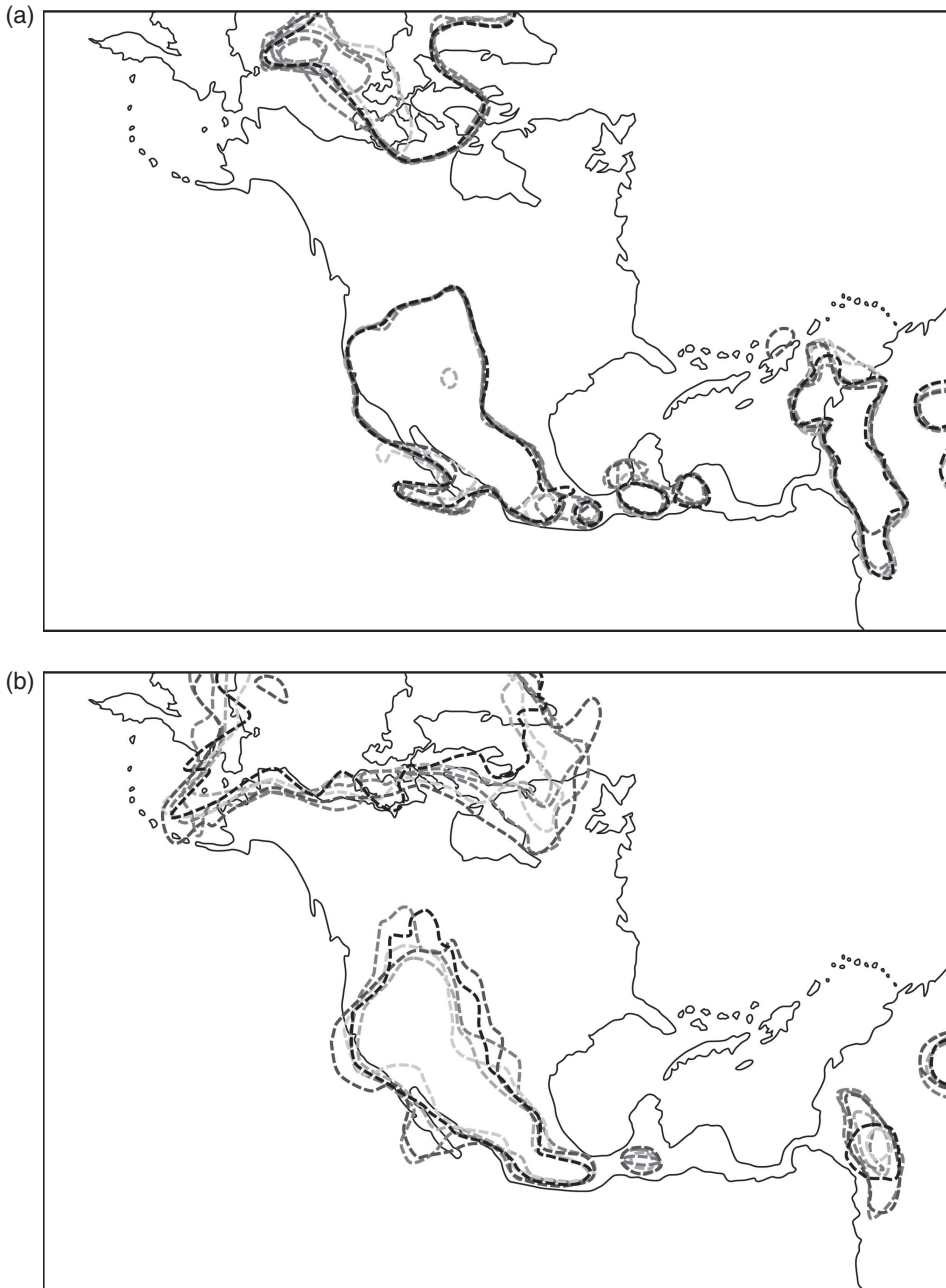
(a)

(b)

Figure 5.6 NCEP spaghetti diagrams of temperature at 850 hPa height, showing model predictions for contours −5°C and 20°C at 00 hours (a) and 192 hours (b). Adapted from the National Weather Service.
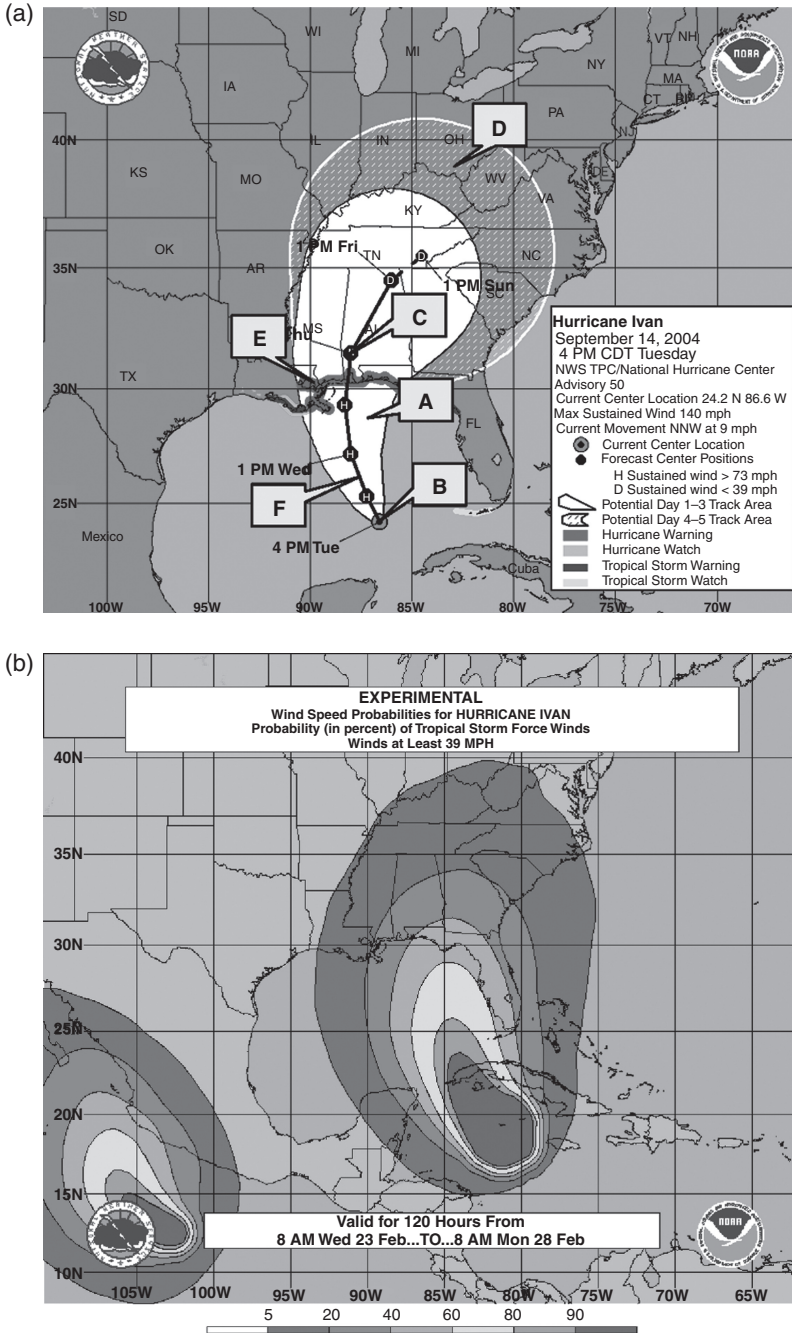
Figure 5.7 The forecast of Hurricane Ivan in 2004 displayed as: (a) The NHC cone of uncertainty (with annotations: see Broad *et al.*, 2007); and (b) an alternative experimental visualisation (Broad *et al.*, 2007) © American Meteorological Society.

Uncertainty estimates are more often provided for seasonal forecasts, reflecting the inherent challenges of predictability. For North Atlantic tropical storms, the UKMO includes uncertainty ranges for the number and ACE up to six months ahead. An example is:

Six tropical storms are predicted as the most likely number to occur in the North Atlantic during the July to November period, with a 70% chance that the number will be in the range three to nine. This represents below-normal activity relative to the 1990–2005 long-term average of 12.4. An ACE index of 60 is predicted as the most likely value, with a 70% chance that the index will be in the range 40 to 80 – which is below normal relative to the 1990–2005 average of 131.

Median, spread and long-term mean are also visualised as barcharts. The simultaneous use of different expressions (language, quantities and graphics) reflects lessons that have been learned about public understanding of risk: it is clear there is no 'one size fits all' to communication or visualisation of uncertainty (Broad *et al.*, 2007; Morss *et al.*, 2008; Joslyn and Nichols, 2009; Joslyn *et al.*, 2009; Spiegelhalter *et al.*, 2011; Stephens *et al.*, 2012).

During a hazard event, effective communication of the emergency response is also essential, to minimise negative impacts (such as lawlessness) that can arise when people are uncertain about their survival, rescue or prospects for evacuation (US House of Representatives, 2006).

5.4.2.2 Emergency response

Most extreme weather disaster events require responses on timescales of hours to days (droughts are an exception). Responses during a hazard event may be proactive (minimising impacts) or reactive (coping with impacts). The former reflect successful implementation of well-designed emergency procedures; the latter reflect poor implementation or poor planning, and may give rise to the need for humanitarian relief (IPCC, 2011). The training exercise Hurricane Pam, for example, was criticised for its 'emphasis on managing the aftermath of the catastrophe and not creating initiatives that would diminish the magnitude' (US House of Representatives, 2006).

Actions include evacuation, search and rescue, providing shelter and giving medical aid, many of which require complex operations including logistics (management of the supply, transport and provision of resources such as blankets and clean water) and effective public communication (Section 5.4.2.1). Hazard response may include aspects that are not essential to immediate survival such as care of mental health (e.g. farmers during drought: Hayes *et al.*, 2004). Police and military personnel may be required to enforce mandatory emergency orders or, in the case of temporary societal breakdown leading to theft or violence, the basic rule of law (e.g. Hurricane Katrina: US House of Representatives, 2006). Response efforts may be hampered as infrastructure can be damaged and services interrupted in unpredictable ways; this is particularly challenging if a disaster occurs on a large spatial scale. The emergency post-disaster phase of Hurricane Katrina was unusually long (emergency sheltering ended only after 14 weeks), 3.5 times longer than the comparable-scale disaster of the 1906 San Francisco earthquake (Kates *et al.*, 2006).

5.4.2.3 Post-impact recovery

Post-event recovery includes restoring infrastructure and services, repairing or replacing damaged buildings and clearing debris. If there are a large number of deaths, the recovery period may also include dealing with corpses. Environmental recovery may include reforestation or reintroduction of wildlife species. Recovery planning should aim to mitigate risk from future disasters (e.g. discouraging rebuilding in risk-prone areas) and if possible make other improvements, such as upgrading services or increasing sustainability (IPCC, 2011). Reddy (2000) examines aspects that make this more or less challenging, such as the level of involvement of stakeholders, in the context of the 1989 Hurricane Hugo. Most aspects of post-disaster recovery are similar to those of other natural hazards. One area in which they differ is that the spatial extent may be very large – an entire country or continent affected by the same hazard – which can hamper recovery of infrastructure and services.

5.4.2.4 Post-impact analysis and lessons learned

A catastrophic event provides a window of opportunity in which it is easier to make policy changes due to increased awareness and availability of resources (CCSP, 2008). Post-event evaluation of the forecasts and decision-making leads to improvements in risk assessment and management: Weisheimer *et al*. (2011), for example, find they can successfully 'hindcast' the 2003 European heat wave once several model parameterisation changes are made. New damage and loss datasets may be used to verify impacts model predictions to improve risk management; however, disaster-related datasets are generally lacking at the local level (IPCC, 2011).

As a result of the severe European heat wave in 2003, the UK introduced the Heat Health Watch System, and during the hot weather of July 2006 only 680 excess deaths were recorded (UK Government Cabinet Office, 2008). Lessons were similarly learned after the Chicago heat wave of 1995 (Klinenberg, 2002). The city had made disastrous mistakes: an emergency was not declared until hundreds of fatalities had occurred; the Fire Department refused paramedic requests for additional staff and ambulances; there was no system to monitor the hospital bypass situation; and the Police Department did not attend to elderly residents. Public relations apparatus was even used to deny there was a disaster and later to define the disaster as natural and unpreventable. When a second heat wave occurred four years later, the city acted quickly and effectively, issuing warnings and press releases to the media, opening cooling centres with free bus transportation, contacting elderly residents and sending police officers and city workers to visit the elderly that lived alone. These actions are thought to have drastically reduced the death toll in the second heat wave: 110 residents died, compared with an excess mortality of 739 four years earlier.

The US House of Representatives report on Hurricane Katrina described some successes in risk assessment and management, including the 'accuracy and timeliness of National Weather Service and National Hurricane Center forecasts' that 'prevented further loss of life'. But it also described a litany of failures and avoidable negative impacts. These included failures in:

(1) *long-term planning*: inadequate levees; inadequately trained and experienced staff; lack of a regional shelter database; an overwhelmed logistics system;

(2) *development of emergency procedures*: lessons not fully learned from the Hurricane Pam exercise; inadequate planning for massive damage to communications; lack of preparation to evacuate and provide medical care for vulnerable groups;

(3) *implementation of emergency procedures*: mandatory evacuation ordered only 19 hours before landfall, rather than at the time of early warning (56 hours); critical elements of the National Response Plan executed late, ineffectively or not at all; poor coordination between federal and state bodies; reactive, rather than proactive, deployment of medical personnel; subjective evacuation decisions for nursing homes that led to preventable deaths; collapse of local law enforcement and lack of effective public communication that led to civil unrest and further delay in relief (US House of Representatives, 2006).

Lessons of risk management may be learned long after the initial recovery period. Homeowners affected by Hurricane Katrina were typically covered by private insurers for wind damage and by the federal government's National Flood Insurance Program (NFIP), or not at all, for flood damage. This led to large uncertainties in forecasting insured losses for insurers, due to the difficulty in predicting damage assessments and disputes (RMS, 2005), and in the extent of coverage for policy-holders due to the cap on NFIP flood claims (Williams, 2008) or lack of flood insurance. Private insurers that issued NFIP flood insurance policies ('write-your-own', WYO) determined the apportioning damages between wind and flooding, which led to conflicts-of-interest, claim disputes and litigation (Williams, 2008). As a result of the flood-related claims of Hurricane Katrina, the NFIP's borrowing authority rose 14-fold to about US$20.8 billion (US GAO, 2007). To ensure the NFIP did not pay too much in future storm events, the US GAO recommended increasing FEMA's access to the procedures and claims information from WYO insurers (Williams, 2008). At the time of writing, these provisions (Flood Insurance Reform Priorities Act of 2011; Flood Insurance Reform Act of 2011) are under consideration by the US Congress.

5.5 Summary

Risk assessment for hydrometeorological hazards builds on a long history of weather forecasting. The definition of these hazards is continuously evolving, from simple univariate metrics to complex multivariate, impacts-related indices, to improve their relevance and usefulness. Nonlinear feedbacks in the climate, and between climate and impacts, make their prediction extremely challenging, and there is a limit of predictability of about a fortnight for atmospheric phenomena. But physical and statistical modelling are strong, diverse research areas that complement each other: the former mainly for event forecasting and warning systems, the latter for longer-term risk management. Both make use of long-term oscillatory modes of natural variability correlated with extreme weather (such as the ENSO) to extend the predictability of risk to many months in advance. Impacts modelling is still catching up with hazard event modelling in terms of comprehensiveness and uncertainty quantification, which is perhaps inevitable given their later position in the causal chain.

Physically based models are generally well-resourced and very frequently tested against observations in operational numerical weather prediction, so their predictive skill is increasing decade-upon-decade: this is due to wider availability and better assimilation of observations, and improvements in models from the addition of physical processes and increases in resolution. Statistical modelling has advanced beyond the early extreme value analyses by incorporating model flexibility, spatial and temporal correlation and non-stationarity, though progress can still be made in these areas and caution must be used in interpreting estimates for the far future. Catastrophe modelling is still catching up with other types of modelling in terms of physical theory and non-stationarity, but leads others in terms of end-to-end assessment of the cascading risks from hazard to financial loss.

In risk management, hydrometeorological hazards are in the fortunate position of having extensive observational and forecasting networks, from longstanding national meteorological agencies and the regional collaborations between them. These give early warnings for extreme weather conditions with varying levels of information content and uncertainty quantification; lessons are still being learned about the most successful approaches. Quantitative decision-making thresholds are incorporated into emergency planning, and policy has long existed for risk mitigation in the form of building codes and other regulation. However, the implementation of risk management strategies has sometimes been lacking, notably during the catastrophic Hurricane Katrina in 2005. Insurers use catastrophe models to assess hurricane risk as a fraction of their capital, which is subject to regulation.

Hydrometeorological hazards have the disadvantage, then, of occurring everywhere under the sun. But longstanding interests in the weather, water availability and ocean conditions have driven the development of sophisticated computational and statistical techniques, and fairly comprehensive observational networks, so that a substantial amount of the risk may be managed.

Acknowledgements

References

AIR Worldwide Corporation (2006) Understanding climatological influences on hurricane activity. *The AIR Near-term Sensitivity Catalog.* pp. 1–16.

Alcamo, J., Acosta-Michlik, L., Carius, A., *et al.* (2008) A new approach to quantifying and comparing vulnerability to drought. *Regional Environmental Change* **8** (4): 137–149.

Alexander, L. V., Zhang, X., Peterson, T. C., *et al.* (2006) Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* **111** (D5): D05109.

Anthes, R. A., Corell, R. W., Holland, G., *et al*. (2006) Hurricanes and global warming: potential linkages and consequences. *Bulletin of the American Meteorological Society* **87** (5): 623–628.

Baccini, M., Biggeri, A., Accetta, G., *et al*. (2008) Heat effects on mortality in 15 European cities. *Epidemiology* **19** (5): 711–719.

Barnett, D., Brown, S., Murphy, J., *et al*. (2006) Quantifying uncertainty in changes in extreme event frequency in response to doubled CO_2 using a large ensemble of GCM simulations. *Climate Dynamics* **26** (5): 489–511.

Barnston, A. G., Mason, S., Goddard, L., *et al*. (2003) Multimodel ensembling in seasonal climate forecasting at IRI. *Bulletin of the American Meteorological Society* **84** (12): 1783–1796.

Barriopedro, D., Fischer, E., Luterbacher, J., *et al*. (2011) The hot summer of 2010: redrawing the temperature record map of Europe. *Science* **332** (6026): 220–224.

Beirlant, J., Goegebeur, Y., Segers, J., *et al*. (2004) *Statistics of Extremes: Theory and Applications*, Hoboken, NJ: John Wiley and Sons.

Beniston, M. (2004) The 2003 heat wave in Europe: a shape of things to come? An analysis based on Swiss climatological data and model simulations. *Geophysical Research Letters* **31** (2): L02202.

Beniston, M., Stephenson, D., Christensen, O., *et al*. (2007) Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change* **81**: 71–95.

Berner, J., Leutbecher, M., Palmer, T., *et al*. (2009) A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences* **66** (3): 603–626.

Berrocal, V. J., Raftery, A. E. and Gneiting, T. (2007) Combining spatial statistical and ensemble information in probabilistic weather forecasts. *Monthly Weather Review* **135** (4): 1386–1402.

Brier, G. (1950) Verification of forecasts expressed in terms of probability. *Monthly Weather Review* **78** (1): 1–3.

Broad, K., Leiserowitz, A., Weinkle, J., *et al*. (2007) Misinterpretations of the 'Cone of Uncertainty' in Florida during the 2004 Hurricane season. *Bulletin of the American Meteorological Society* **88** (5): 651–667.

Brohan, P., Kennedy, J., Harris, S., *et al*. (2006) Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. *Journal of Geophysical Research-Atmospheres* **111** (D12): D12106.

Brown, S. J., Caesar, J. and Ferro, C. A. T. (2008) Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research-Atmospheres* **113** (D5): D05115.

Browning, K. A. (2004) The sting at the end of the tail: damaging winds associated with extratropical cyclones. *Quarterly Journal of the Royal Meteorological Society* **130** (597): 375–399.

Bründl, M. and Rickli, C. (2002), The storm Lothar 1999 in Switzerland: an incident analysis. *Forest, Snow and Landscape Research*, **77**(1/2): 207–217.

Buizza, R., Miller, M. and Palmer, T. (1999) Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society* **125** (560): 2887–2908.

Burke, E. J. and Brown, S. J. (2008) Evaluating uncertainties in the projection of future drought. *Journal of Hydrometeorology* **9** (2): 292–299.

Camargo, S. J. and Barnston, A. G. (2009) Experimental dynamical seasonal forecasts of tropical cyclone activity at IRI. *Weather and Forecasting* **24** (2): 472–491.

Camargo, S. J., Barnston, A., Klotzbach, P., *et al.* (2007) Seasonal tropical cyclone forecasts. *WMO Bulletin* **56** (4): 297–309.

Casati, B., Wilson, L. and Stephenson, D. (2008) Forecast verification: current status and future directions. *Meteorological Applications* **15**: 3–18.

CCSP (2008) *Weather and Climate Extremes in a Changing Climate. Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands.* Washington, DC: Department of Commerce, NOAA's National Climatic Data Center.

Chatfield, C. (2004) *The Analysis of Time Series: An Introduction*, 6th edn, Boca Raton, FL: CRC Press.

Chavez-Demoulin, V. and Davison, A. (2005) Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society Series C-Applied Statistics* **54** (1): 207–222.

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*, New York, NY: Springer.

Coles, S., Pericchi, L. and Sisson, S. (2003) A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology* **273** (1–4): 35–50.

Cummins, J. (2007) Reinsurance for natural and man-made catastrophes in the United States: current state of the market and regulatory reforms. *Risk Management and Insurance Review* **10** (2): 179–220.

Curriero, F., Patz, J., Rose, J., *et al.* (2001) The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *American Journal of Public Health* **91** (8): 1194.

Dai, A. (2011) Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change* **2**: 45–65.

Davison, A. C. and Hinkley, D. V. (1997) *Bootstrap Methods and their Application*, Cambridge: Cambridge University Press.

De Bono, A., Peduzzi, P. and Giuliani, G. (2004) Impacts of summer 2003 heat wave in Europe. *Early Warning on Emerging Environmental Threats* **2**: n.p.

Della-Marta, P. M., Mathis, H., Frei, C., *et al.* (2009) The return period of wind storms over Europe. *International Journal of Climatology* **29** (3): 437–459.

Dlugolecki, A., *et al.* (2009) Climate change and its implications for catastrophe modelling. In *Coping with Climate Change: Risks and Opportunites for Insurers*, London: Chartered Insurance Institute, pp. 1–23.

Dorigo, W. A., Wagner, W., Hohensinn, R., *et al.* (2011) The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences Discussions* **8** (1): 1609–1663.

Dysthe, K., Krogstad, H. E. and Muller, P. (2008) Oceanic rogue waves. *Annual Review of Fluid Mechanics* **40**: 287–310.

Eastoe, E. and Tawn, J. (2009) Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society Series C-Applied Statistics* **58** (1): 25–45.

Edwards, A., Elwyn, G. and Mulley, A. (2002) Explaining risks: turning numerical data into meaningful pictures. *British Medical Journal* **324** (7341): 827–830.

El Adlouni, S., Ouarda, T., Zhang, X., *et al.* (2007) Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resources Research* **43** (3): W03410.

Elsner, J. (2003) Tracking hurricanes. *Bulletin of the American Meteorological Society* **84** (3): 353–356.

Elsner, J. and Bossak, B. H. (2004) Hurricane Landfall Probability and climate. In *Hurricanes and Typhoons: Past, Present and Future*, New York, NY: Columbia University Press, pp. 333–353.

Elsner, J. and Jagger, T. H. (2006) Prediction models for annual US hurricane counts. *Journal of Climate* **19** (12): 2935–2952.

Elsner, J., Niu, X. and Jagger, T. (2004) Detecting shifts in hurricane rates using a Markov chain Monte Carlo approach. *Journal of Climate* **17** (13): 2652–2666.

Emanuel, K. (2005) Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **436** (7051): 686–688.

Emanuel, K., Ravela, S., Vivant, E., *et al*. (2006) A statistical deterministic approach to hurricane risk assessment. *Bulletin of the American Meteorological Society* **87** (3): 299–314.

Engeland, K., Hisdal, H. and Frigessi, A. (2004) Practical extreme value modelling of hydrological floods and droughts: a case study. *Extremes* **7**: 5–30.

Evensen, G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research* **99** (C5): 10143–10162.

Evensen, G. (2009) *Data Assimilation: The Ensemble Kalman Filter*, New York, NY: Springer.

Fairchild, A., Colgrove, J. and Moser Jones, Marian (2006) The challenge of mandatory evacuation: providing for and deciding for. *Health Affairs* **25** (4): 958–967.

FEMA (2007) HAZUS-MH MR3 Technical Manual.

Ferro, C. (2007) A probability model for verifying deterministic forecasts of extreme events. *Weather and Forecasting* **22** (5): 1089–1100.

Ferro, C. A. T., Richardson, D. S. and Weigel, A. P. (2008) On the effect of ensemble size on the discrete and continuous ranked probability scores. *Meteorological Applications* **15** (1): 19–24.

Fink, A. H., Brücher, B., Ermert, V., *et al*. (2009) The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Natural Hazards and Earth System Science* **9**: 405–423.

Forristall (2005) Understanding rogue waves: Are new physics really necessary? Proceedings of the 14th 'Aha Huliko'a Winter Workshop 2005 on Rogue Waves, 25–28 January, Honolulu, USA, pp. 1–7.

Frich, P., Alexander, L., Della-Martin, P., *et al*. (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research* **19** (3): 193–212.

Funk, C. (2011) We thought trouble was coming. *Nature* **476** (7358): 7.

Githeko, A. and Ndegwa, W. (2001) Predicting malaria epidemics in the Kenyan highlands using climate data: a tool for decision makers. *Global Change & Human Health* **2** (1): 54–63.

Gleick, P. H., Cooley, H. and Katz, D. (2006) *The World's Water, 2006–2007: The Biennial Report on Freshwater Resources*, Washington, DC: Island Press.

Gneiting, T. (2005) Atmospheric science: weather forecasting with ensemble methods. *Science* **310** (5746): 248–249.

Gosling, S. N., Lowe, S., McGregor, G., *et al*. (2009) Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change* **92** (3–4): 299–341.

Grossi, P. and Kunreuther, H. (eds) (2005) *Catastrophe Modeling: A New Approach to Managing Risk*, New York, NY: Springer.

Hagedorn, R., Doblas-Reyes, F. J. and Palmer, T. N. (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting: I – basic concept. *Tellus Series A – Dynamic Meteorology and Oceanography* **57** (3): 219–233.

Hall, T. M. and Jewson, S. (2007) Statistical modelling of North Atlantic tropical cyclone tracks. *Tellus Series A-Dynamic Meteorology and Oceanography* **59** (4): 486–498.

Hall, T. M. and Jewson, S. (2008) Comparison of local and basinwide methods for risk assessment of tropical cyclone landfall. *Journal of Applied Meteorology and Climatology* **47** (2): 361–367.

Hassan, R. M., Scholes, R. and Ash, N. (eds) (2005) *Ecosystems and Human Well-being: Current State and Trends – Findings of the Condition and Trends Working Group of the Millennium Ecosystem Assessment*, Washington,. DC: Island Press.

Hayes, M., Wilhelmi, O. and Knutson, C. (2004) Reducing drought risk: bridging theory and practice. *Natural Hazards Review* **5**: 106.

Held, I. M. and Soden, B. J. (2006) Robust responses of the hydrological cycle to global warming. *Journal of Climate* **19** (21): 5686–5699.

Hirschi, M., Seneviratne, S., Orlowsky, B., *et al*. (2011) Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nature Geoscience* **4** (1): 17–21.

Holliday, N. P., Yelland, M., Pascal, R., *et al*. (2006) Were extreme waves in the Rockall Trough the largest ever recorded? *Geophysical Research Letters* **33** (5): L05613.

Hollingsworth, A. and Lönnberg, P. (1986) The statistical structure of short-range forecast errors as determined from radiosonde data. Part 1: the wind field. *Tellus Series A – Dynamic Meteorology and Oceanography* **38** (2): 111–136.

Houtekamer, P., Mitchel, H., Pellerin, G., *et al*. (2005) Atmospheric data assimilation with an ensemble Kalman filter: results with real observations. *Monthly Weather Review* **133** (3): 604–620.

Houtekamer, P. L., Mitchell, H. L. and Deng, X. (2009) Model error representation in an operational ensemble Kalman filter. *Monthly Weather Review* **137** (7): 2126–2143.

Iman, R., Johnson, M. and Watson, C. (2005a) Uncertainty analysis for computer model projections of hurricane losses. *Risk Analysis* **25** (5): 1299–1312.

Iman, R., Johnson, M. E. and Watson, C. C. (2005b) Sensitivity analysis for computer model projections of hurricane losses. *Risk Analysis* **25** (5): 1277–1297.

IPCC (2007a) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.

IPCC (2007b) *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.

IPCC (2011) *Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX) Summary for Policymakers*, Geneva: IPCC.

Jagger, T. H. and Elsner, J. B. (2006) Climatology models for extreme hurricane winds near the United States. *Journal of Climate* **19** (13): 3220–3236.

Joacim, R., Kristie, E. and Bertil, F. (2010) Mortality related to temperature and persistent extreme temperatures: a study of cause-specific and age-stratified mortality. *Occupational and Environmental Medicine* **68** (7): 531–536.

Jolliffe, I. T. and Stephenson, D. B. (2003) *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, Chichester: John Wiley and Sons.

Jones, P. D., *et al*. (2009) Projections of future daily climate for the UK from the Weather Generator. In *UK Climate Projections*, Exeter: Met Office Hadley Centre, pp. 1–48.

Joslyn, S. L. and Nichols, R. M. (2009) Probability or frequency? Expressing forecast uncertainty in public weather forecasts. *Meteorological Applications* **16** (3): 309–314.

Joslyn, S. and Savelli, S. (2010) Communicating forecast uncertainty: public perception of weather forecast uncertainty. *Meteorological Applications* **17** (2): 180–195.

Joslyn, S., Nadav-Greenberg, L. and Nichols, R. M. (2009) Probability of precipitation assessment and enhancement of end-user understanding. *Bulletin of the American Meteorological Society* **90** (2): 185–193.

Kallis, G. (2008) Droughts. *Annual Review of Environment And Resources* **33**: 85–118.

Kalnay, E. (2003) *Atmospheric Modeling, Data Assimilation, and Predictability*, Cambridge: Cambridge University Press.

Kates, R. W., Colten, C. E., Laska, S., *et al*. (2006) Reconstruction of New Orleans after Hurricane Katrina: a research perspective. *Proceedings of the National Academy of Sciences* **103** (40): 14653–14660.

Khaliq, M. N., Ouarda, T., Ondo, J., *et al*. (2006) Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: a review. *Journal of Hydrology* **329** (3–4): 534–552.

Kharif, C. and Pelinovsky, E. (2003) Physical mechanisms of the rogue wave phenomenon. *European Journal of Mechanics B – Fluids* **22** (6): 603–634.

Kharif, C., Pelinovsky, E. and Slunyaev, A. (2009) *Rogue Waves in the Ocean*, New York, NY: Springer.

Kleiber, W., Raftery, A., Baars, J., *et al*. (2010) Locally calibrated probabilistic temperature forecasting using geostatistical model averaging and local Bayesian model averaging. *Monthly Weather Review* **139** (8): 2630–2649.

Klinenberg, E. (2002) *Heat Wave: A Social Autopsy of Disaster in Chicago*, Chicago, IL: University of Chicago Press.

Klotzbach, P., Barnston, A. G., Bell, G., *et al*. (2011) Seasonal forecasting of tropical cyclones. In *Global Guide to Tropical Cyclone Forecasting*, ed. C. Guard, Geneva: World Meteorological Organization.

Kovats, R. S. and Hajat, S. (2008) Heat stress and public health: a critical review. *Annual Review of Public Health* **29** (1): 41–55.

Kundzewicz, Z. and Robson, A. (2004) Change detection in hydrological records: a review of the methodology. *Hydrological Sciences Journal* **49** (1): 7–19.

Kunkel, K., Andsager, K. and Easterling, D. (1999) Long-term trends in extreme precipitation events over the conterminous United States and Canada. *Journal of Climate* **12** (8): 2515–2527.

Kyselý, J. (2008) A cautionary note on the use of nonparametric bootstrap for estimating uncertainties in extreme-value models. *Journal of Applied Meteorology and Climatology* **47** (12): 3236–3251.

Kyselý, J. (2009) Recent severe heat waves in central Europe: how to view them in a long-term prospect? *International Journal of Climatology* **30** (1): 89–109.

Kyselý, J. and Picek, J. (2007) Regional growth curves and improved design value estimates of extreme precipitation events in the Czech Republic. *Climate Research* **33** (3): 243–255.

Landsea, C. (2007) Counting Atlantic tropical cyclones back to 1900. *Eos* **88** (18): 197–208.

Landsea, C., Harper, B., Hoarau, K., *et al*. (2006) Can we detect trends in extreme tropical cyclones? *Science* **313** (5786): 452–454.

Lawton, G. (2001) Monsters of the deep. *New Scientist* **170** (2297): 28–32. http://www.newscientist.com/article/mg17022974.900-monsters-of-the-deep.html.

Leary, N., Conde, C. and Kulkarni, J., *et al*. (ed.) (2008) *Climate Change and Vulnerability*, London: Earthscan.

Lemeshko, B., Lemeshko, S. and Postovalov, S. (2009) Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses: I. The verification of simple hypotheses. *Journal of Applied and Industrial Mathematics* **3** (4): 462–475.

Lemeshko, B., Lemeshko, S. and Postovalov, S. (2010) Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses: II. Verification of complex hypotheses. *Journal of Applied and Industrial Mathematics* **4** (1): 79–93.

Levin, Z. and Cotton, W. R. (eds) (2008) *Aerosol Pollution Impact on Precipitation: A Scientific Review*, New York, NY: Springer-Science Press.

Li, Y., Cai, W. and Campbell, E. (2005) Statistical modeling of extreme rainfall in southwest Western Australia. *Journal of Climate* **18** (6): 852–863.

Li, Z., Niu, F., Fan, J., *et al*. (2011) Long-term impacts of aerosols on the vertical development of clouds and precipitation. *Nature Geoscience* **4** (12): 888–894.

Lloyd-Hughes, B. and Saunders, M. (2002) A drought climatology for Europe. *International Journal of Climatology* **22** (13): 1571–1592.

Lorenz, E. N. (1965) A study of the predictability of a 28-variable atmospheric model. *Tellus* **17**: 321–333.

Lynch, P. (2008) The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, **227** (7): 3431–3444.

Mailier, P. J., Stephenson, D., Ferro, C., *et al*. (2006) Serial clustering of extratropical cyclones. *Monthly Weather Review* **134** (8): 2224–2240.

McFarlane, N. (2011) Parameterizations: representing key processes in climate models without resolving them. *Wiley Interdisciplinary Reviews: Climate Change* **2** (4): 482–497.

McGregor, P. and MacDougall, K. (2009) A review of the Scottish rain-gauge network. *Proceedings of the Institution of Civil Engineers – Water Management* **162** (2): 137–146.

McNeil, A. J., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton, NJ: Princeton University Press.

Mehrotra, R., Srikanthan, R. and Sharma, A. (2006) A comparison of three stochastic multi-site precipitation occurrence generators. *Journal of Hydrology* **331**: 280–292.

Menéndez, M., Mendez, F., Izahuirre, C., *et al*. (2009) The influence of seasonality on estimating return values of significant wave height. *Coastal Engineering* **56** (3): 211–219.

Michelozzi, P., Kirchmayer, U., Katsouyanni, K., *et al*. (2007) Assessment and prevention of acute health effects of weather conditions in Europe, the PHEWE project: background, objectives, design. *Environmental Health* **6**: 12.

Miller, S., Muir-Wood, R. and Boissonade, A. (2008) An exploration of trends in normalized weather-related catastrophe losses. In *Climate Extremes and Society*, ed. H. F. Diaz and R. J. Murnane, Cambridge: Cambridge Universiy Press, pp. 225–247.

Morss, R. E., Demuth, J. L. and Lazo, J. K. (2008) Communicating uncertainty in weather forecasts: a survey of the U.S. public. *Weather and Forecasting* **23** (5): 974–991.

Nelsen, R. (2006) *An Introduction to Copulas*, 2nd edn, New York, NY: Springer.

Nicolosi, V., Cancelliere, A. and Rossi, G. (2009) Reducing risk of shortages due to drought in water supply systems using genetic algorithms. *Irrigation and Drainage* **58** (2): 171–188.

Palmer, T. (2001) A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society* **127** (572): 279–304.

Palmer, T., Alessandri, A., Andersen, U., *et al*. (2004) Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society* **85** (6): 853–872.

Palmer, T., Shutts, G., Hagedorn, F., *et al*. (2005) Representing model uncertainty in weather and climate prediction. *Annual Review of Earth and Planetary Sciences* **33**: 163–193.

Palmer, W. C. (1965) *Meteorological Drought*, Washington, DC: US Department of Commerce Weather Bureau.

Pang, W., Forster, J. and Troutt, M. (2001) Estimation of wind speed distribution using Markov chain Monte Carlo techniques. *Journal of Applied Meterology* **40**: 1476–1484.

Park, J. and Jung, H. (2002) Modelling Korean extreme rainfall using a Kappa distribution and maximum likelihood estimate. *Theoretical and Applied Climatology* **72** (1): 55–64.

Parker, D., Legg, T. and Folland, C. (1992) A new daily central England temperature series, 1772–1991. *International Journal of Climatology* **12** (4): 317–342.

Parrish, D. and Derber, J. (1992) The National-Meteorological-Center's spectral statistical-interpolation analysis system. *Monthly Weather Review* **120** (8): 1747–1763.

Persson, A. and Grazzini, F. (2007) *User Guide to ECMWF Forecast Products*, 4th edn, Reading: ECMWF.

Pielke, R. (2005) Are there trends in hurricane destruction? *Nature* **438** (7071): E11.

Pielke, R., Landsea, C., Mayfiled, M., *et al*. (2006) Reply to 'Hurricanes and global warming – Potential linkages and consequences'. *Bulletin of the American Meteorological Society* **87** (5): 628–631.

Pielke, R., Gratz, J., Landsea, C., *et al*. (2008) Normalized hurricane damage in the United States: 1900–2005. *Natural Hazards Review* **9**: 29.

Pinelli, J., Gurley, K. and Subramanian, C. (2008) Validation of a probabilistic model for hurricane insurance loss projections in Florida. *Reliability Engineering and System Safety* **93**: 1896–1905.

Powell, M., Soukup, G., Gulati, S., *et al*. (2005) State of Florida hurricane loss projection model: atmospheric science component. *Journal of Wind Engineering and Industrial Aerodynamics* **93** (8): 651–674.

Rabier, F., Järvinen, H.s and Klinker, E. (2000) The ECMWF operational implementation of four-dimensional variational assimilation: I. Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society* **126**: 1143–1170.

Raftery, A., Balabdaoui, F., Gneiting, T., *et al*. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review* **133** (5): 1155–1174.

Ramesh, N. and Davison, A. (2002) Local models for exploratory analysis of hydrological extremes. *Journal of Hydrology* **256**: 106–119.

Reddy, S. (2000) Factors influencing the incorporation of hazard mitigation during recovery from disaster. *Natural Hazards* **22** (2): 185–201.

Ribereau, P., Guillou, A. and Naveau, P. (2008) Estimating return levels from maxima of non-stationary random sequences using the Generalized PWM method. *Nonlinear Processes in Geophysics* **15** (6): 1033–1039.

Rivière, G. and Orlanski, I. (2007) Characteristics of the Atlantic storm-track eddy activity and its relation with the North Atlantic oscillation. *Journal of the Atmospheric Sciences* **64** (2): 241–266.

RMS (2005) *Hurricane Katrina: Lessons and Implications for Catastrophe Risk Management*, Newark, CA: RMS.

RMS (2008) *A Guide to Catastrophe Modeling*, Newark, CA: RMS.

Robeson, S. M. (2008) Applied climatology: drought. *Progress in Physical Geography* **32** (3): 303–309.

Robine, J., Cheung, S., Le Roy, S., et al. (2008) Death toll exceeded 70,000 in Europe during the summer of 2003. *Comptes Rendus Biologies* **331** (2): 171–178.

Robinson, P. (2001) On the definition of a heat wave. *Journal of Applied Meteorology* **40** (4): 762–775.

Roulston, M. S. and Kaplan, T. R. (2009) A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. *Meteorological Applications* **16** (2): 237–244.

Roulston, M., Bolton, G., Kleit, A., et al. (2006) A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting* **21** (1): 116–122.

Rusticucci, M. and Vargas, W. (2002) Cold and warm events over Argentina and their relationship with the ENSO phases: risk evaluation analysis. *International Journal of Climatology* **22** (4): 467–483.

Salvadori, G., Kottegoda, N., De Michele, C., et al. (2007) *Extremes in Nature: An Approach Using Copulas*, Berlin: Springer.

Sanders, F. and Gyakum, J. R. (1980) Synoptic-dynamic climatology of the 'bomb.' *Monthly Weather Review* **108** (10): 1589–1606.

Saunders, M. A. and Lea, A. S. (2005) Seasonal prediction of hurricane activity reaching the coast of the United States. *Nature* **434** (7036): 1005–1008.

Schär, C., Vidale, P., Lüthi, D., et al. (2004) The role of increasing temperature variability in European summer heatwaves. *Nature* **427** (6972): 332–336.

Schmidt, S., Kemfert, C. and Höppe, P. (2009) Tropical cyclone losses in the USA and the impact of climate change: a trend analysis based on data from a new approach to adjusting storm losses. *Environmental Impact Assessment Review* **29** (6): 359–369.

Schölzel, C. and Friederichs, P. (2008) Multivariate non-normally distributed random variables in climate research: introduction to the copula approach. *Nonlinear Processes in Geophysics* **15** (5): 761–772.

Senay, G. and Verdin, J. (2003) Characterization of yield reduction in Ethiopia using a GIS-based crop water balance model. *Canadian Journal of Remote Sensing* **29** (6): 687–692.

Sheridan, S. (2002) The redevelopment of a weather-type classification scheme for North America. *International Journal of Climatology* **22** (1): 51–68.

Sheridan, S. and Kalkstein, L. (2004) Progress in heat watch-warning system technology. *Bulletin of the American Meteorological Society* **85** (12): 1931–1941.

Shutts, G. (2005) A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society* **131** (612): 3079–3102.

Smoyer-Tomic, K., Kuhn, R. and Hudson, A. (2003) Heat wave hazards: an overview of heat wave impacts in Canada. *Natural Hazards* **28** (2–3): 463–485.

Souch, C. and Grimmond, C. (2004) Applied climatology: 'heat waves.' *Progress in Physical Geography* **28** (4): 599–606.

Spiegelhalter, D., Pearson, M. and Short, I. (2011) Visualizing uncertainty about the future. *Science* **333** (6048): 1393–1400.

Stephens, G. L., L'Ecuyer, T., Suzuki, K., *et al*. (2010) Dreary state of precipitation in global models. *Journal of Geophysical Research* **115** (D24): D24211.

Stephens, E. M., Edwards, T. L. and Demeritt, D. (2012) Communicating probabilistic information from climate model ensembles – lessons from numerical weather prediction. *Wiley Interdisciplinary Reviews: Climate Change* **3** (5): 409–426.

Stephenson, D., Casati, B., Ferro, C. A. T., *et al*. (2008) The extreme dependency score: a non-vanishing measure for forecasts of rare events. *Meteorological Applications* **15** (1): 41–50.

Storch, H. von (1995) Misuses of statistical analysis in climate research. In *Analysis of Climate Variability: Applications of Statistical Techniques*, ed., H. von Storch and A. Navarra, New York, NY: Springer Verlag, pp. 11–26.

Storch, H. von and Zwiers, F. W. (2003) *Statistical Analysis in Climate Research*, Cambridge: Cambridge University Press.

Talagrand, O., Vautard, R. and Strauss, B. (1997) Evaluation of probabilistic prediction systems. In *ECMWF Workshop on Predictability*, Reading: ECMWF.

Tebaldi, C., Hayhoe, K., Arblaster, J., *et al*. (2006) Going to the extremes. *Climatic Change* **79** (3–4): 185–211.

Tolman, H. (2009) User manual and system documentation of WAVEWATCH III TM version 3.14, NOAA/NWS/NCEP/MMAB Technical Note 276.

Townsend, F. F. (2006) *The Federal Response to Hurricane Katrina: Lessons Learned*, Washington DC: The White House.

Tucker, M. J. and Pitt, E. G. (eds) (2001) *Waves in Ocean Engineering*, New York, NY: Elsevier.

UK Environment Agency (2009) Drought plan for Anglian region.

UK Government Cabinet Office (2008) National risk register. http://www.cabinetoffice.gov.uk.

United Nations (2006) *Global Survey of Early Warning Systems: An Assessment of Capacities, Gaps and Opportunities Towards Building a Comprehensive Global Early Warning System For All Natural Hazards*, Geneva: United Nations.

Uppala, S. M., Kållberg, P., Simmons, A., *et al*. (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131** (612): 2961–3012.

USDA (1994) *Major World Crop Areas and Climatic Profiles*, Washington, DC: World Agricultural Outlook Board, U.S. Department of Agriculture.

US GAO (2007) GAO-07-169 National Flood Insurance Program: New Processes Aided Hurricane Katrina Claims Handling, but FEMA's Oversight Should Be Improved. pp. 1–68.

US House of Representatives (2006) A failure of initiative: final report of the Select Bipartisan Committee to Investigate the Preparation for and Response to Hurricane Katrina.

van Leeuwen, P. (2009) Particle filtering in geophysical systems. *American Meterological Society* **137**: 4089–4114.

van Leeuwen, P. (2010) Nonlinear data assimilation in geosciences: an extremely efficient particle filter. *Quarterly Journal of the Royal Meteorological Society* **136** (653): 1991–1999.

van Leeuwen, P. (2011) Efficient nonlinear data-assimilation in geophysical fluid dynamics. *Computers & Fluids* **46** (1): 52–58.

Vecchi, G. A. and Knutson, T. R. (2011) Estimating annual numbers of Atlantic hurricanes missing from the HURDAT database (1878–1965) Using Ship Track Density. *American Meteorological Society* **24** (4): 1736–1746.

Vickery, P., Skerlj, P., and Steckley, A., *et al*. (2000a) Hurricane wind field model for use in hurricane simulations. *Journal of Structural Engineering* **126** (10): 1203–1221.

Vickery, P., Skerlj, P. and Twisdale, L.s (2000b) Simulation of hurricane risk in the US using empirical track model. *Journal of Structural Engineering* **126** (10): 1222–1237.

Vickery, P., Lin, J., Skerlj, P., *et al.* (2006a) HAZUS-MH hurricane model methodology: I. Hurricane hazard, terrain, and wind load modeling. *Natural Hazards Review* **7**: 82–92.

Vickery, P., Skerlj, P., Lin, J., *et al.* (2006b) HAZUS-MH hurricane model methodology: II. Damage and loss estimation. *Natural Hazards* **7**: 94–103.

Vitart, Frédéric (2006) Seasonal forecasting of tropical storm frequency using a multi-model ensemble. *Quarterly Journal of the Royal Meteorological Society* **132** (615): 647–666.

Vitart, F. D. and Stockdale, T. (2001) Seasonal forecasting of tropical storms using coupled GCM integrations. *Monthly Weather Review* **129** (10): 2521–2537.

Vitart, F., Huddleston, M., Déqué, M., *et al.* (2007) Dynamically-based seasonal forecasts of Atlantic tropical storm activity issued in June by EUROSIP. *Geophysical Research Letters* **34** (16): L16815.

Vitolo, R., Stephenson, D., Cook, I., *et al.* (2009) Serial clustering of intense European storms. *Meteorologische Zeitschrift* **18** (4): 411–424.

Webster, P., Holland, G., Curry, J., *et al.* (2005) Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science* **309** (5742): 1844.

Weisheimer, A., Doblas-Reyes, F., Palmer. T., *et al.* (2009) ENSEMBLES: a new multi-model ensemble for seasonal-to-annual predictions – skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophysical Research Letters* **36**: L21711.

Weisheimer, A., Doblas-Reyes, F., Jung, T., *et al.* (2011) On the predictability of the extreme summer 2003 over Europe. *Geophysical Research Letters* **38** (5): L05704.

Wikle, C. and Berliner, L. (2007) A Bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena* **230** (1–2): 1–16.

Wilks, D. S. (2011) *Statistical Methods in the Atmospheric Sciences: An Introduction*, 3rd edn, Oxford: Academic Press.

Wilks, D. S. and Wilby, R. L. (1999) The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23** (3): 329–357.

Williams, O. (2008) *National Flood Insurance Program: Greater Transparency and Oversight of Wind and Flood Damage Determinations are Needed*, Washington, DC: US GAO.

Wimmer, W., Challenor, P. and Retzler, C. (2006) Extreme wave heights in the north Atlantic from altimeter data. *Renewable Energy* **31** (2): 241–248.

Woolf, D., Challenor, P. and Cotton, P. (2002) Variability and predictability of the North Atlantic wave climate. *Journal of Geophysical Research – Oceans* **107**: 3145.

Wu, H., Hayes, M., Weiss, A., *et al.* (2001) An evaluation of the Standardized Precipitation Index, the China-Z Index and the statistical Z-Score. *International Journal Of Climatology* **21** (6): 745–758.

Young, I. R. (1999) Seasonal variability of the global ocean wind and wave climate. *International Journal of Climatology* **19** (9): 931–950.

Zhang, X., Hogg, W. and Mekis, E. (2001) Spatial and temporal characteristics of heavy precipitation events over Canada. *Journal of Climate* **14** (9): 1923–1936.

Zhao, X. and Chu, P.-S. (2010) Bayesian changepoint analysis for extreme events (typhoons, heavy rainfall, and heat waves): an RJMCMC approach. *Journal of Climate* **23** (5): 1034–1046.

6

Hydrometeorological hazards under future climate change

T. L. EDWARDS

6.1 Introduction

The climate is always changing. Natural causes, such as volcano eruptions and variations in solar output, are continuously nudging the radiation balance of the atmosphere, causing it to warm or cool, and the oceans, land surface and cryosphere respond. Since the Industrial Revolution (IPCC, 2007a), perhaps much earlier (Ruddiman, 2003), humans have also been affecting climate by changing the surface of the land and the composition of the atmosphere. The net effect of these changes over the twentieth century has been a warming of the climate (IPCC, 2007a).

Climate change is a change in the statistical properties of weather, usually defined on timescales of three decades or more: not only in the mean ('what we expect', coined by science fiction writer Robert A. Heinlein in 1973), but the full distribution. Weather hazards lie in the extreme tails of this distribution. As mean global temperature increases, the distribution shifts towards some extreme weather hazards such as heat waves, and if it also changes shape (widens or skews) then the magnitude of these extremes could change more quickly than the mean. So far the picture appears to be mixed, with maximum temperatures changing more quickly than the mean since 1950 in some areas and more slowly in others (Brown *et al.*, 2008). Further increases in global mean temperature are predicted to occur this century (based on inertia in the climate system, inertia in societal change and a range of plausible long-term changes), and several studies have predicted that the intensity of extreme high temperatures and precipitation are likely to increase more rapidly than the mean in many areas (e.g. Beniston *et al.*, 2007; Kharin *et al.*, 2007). Predictions for extreme weather events are summarised by the Intergovernmental Panel on Climate Change (IPCC) in the recent *Special Report on Managing the Risks of Extreme Events and Disasters* (IPCC, 2012).

Risk assessment and management for present-day hydrometeorological hazards are described in the previous chapter; for convenience some of the main points are summarised here. Extreme weather is often defined (Section 5.2.1) in terms of exceedances of a threshold, often multivariate (i.e. extremes of multiple variables), where the threshold is either fixed or relative to local climatology. Some alternative definitions are the highest value(s) of a fixed time period, such as annual maxima, or definitions based on synoptic classifications.

Extreme weather events have no distinct trigger but may be associated with persistent atmospheric states, such as blocking, or particular modes of natural variability, such as the El Niño Southern Oscillation (ENSO: Section 5.2.2). Forecasts are made with both physically based, computationally expensive computer models (Section 5.3.2) and statistical methods (Section 5.3.3), for lead times ranging from days to seasons. These are necessarily probabilistic, due to the chaotic nature of the atmosphere. For physically based models, ensemble forecasts of hazard frequencies are generated by perturbing initial conditions and other uncertain quantities. In statistical methods, probabilistic forecasts are generated by inferring a statistical relationship between the hazard and a persistent atmosphere–ocean state. Long-term return values are also estimated, typically extrapolating to rarer extremes and mostly under an assumption of stationarity. Both physically based model ensembles and statistical models are calibrated using historical and recent observations. Uncertainty assessment relies upon the quality and stationarity of the observational record, and our ability to quantify the limitations of the physical or statistical model.

Exposure to extreme weather is global, though exposure to particular hazard types depends on local factors such as topography, proximity to the coast and hydrological feedbacks; losses occur to all sectors, including human health, infrastructure and business and ecosystems (Section 5.2.4). Loss modelling is typically statistical, because the physical mechanisms are often not well understood. Catastrophe modelling (Section 5.3.3), both proprietary and open, incorporates the full causal chain from hazard to loss using various statistical modules, though most are limited in their ability to incorporate physical plausibility constraints. Risk management strategies for extreme weather encompass hazard-monitoring networks (weather forecasting and observational networks) that trigger real-time alerts and actions and long-term planning to reduce, for example, vulnerability of housing or insured assets (Section 5.4.1). Hazard communication and response are challenging for extreme weather events, because they are multivariate, probabilistic and depend on interactions at small spatial scales while affecting large regions; lessons are inevitably learned in post-event recovery and analysis (Section 5.4.2).

Extreme weather hazards have very serious consequences today, so assessment and management of the potential future changes in their risks is therefore of great importance to society. This chapter summarises the risks and uncertainties of droughts, heat waves, extreme precipitation and wind storms under future climate change. Freak ocean waves (Chapter 5) are not discussed. Attribution of individual extreme weather events to human causes is a rapidly developing research area (e.g. Stott *et al.*, 2004; IPCC, 2012; Min *et al.*, 2011; Pall *et al.*, 2011), and not covered here. Section 6.2 discusses methods of risk assessment for hazards and their impacts; Section 6.3 describes risk management and communication; and Section 6.4 summarises.

6.2 Risk assessment

This section describes the challenges and current methods of assessing risk for extreme weather hazards under future climate change. In climate science, the term 'risk' has often

been used in the sense of 'probability of occurrence' of, for example, a particular hazard event. But true risk assessments, which also incorporate hazard impacts, are becoming more common; this is the sense in which the word 'risk' is used in this and the previous chapter (a more detailed definition is given in Chapter 2). Risk assessment for weather hazards under future climate change encompasses multiple research areas in a long causal chain. First, there is consideration of the range of possible future driving factors, both human and natural, that influence climate. Second, there is prediction of the future response of climate to those factors, including not only mean climate but the extremes. Third, there is assessment of the future impacts of these hazards, including changes in exposure and vulnerability. Uncertainties must be propagated along this causal chain.

The main differences between extreme weather in the present day (Chapter 5) and under future climate change (this chapter) are: the long timescale on which climate is defined; the resulting importance of changing boundary conditions; and the unpredictability of long-term changes in those boundary conditions. Boundary conditions are the external controls of the system evolution, such as atmospheric composition (e.g. greenhouse gas concentrations), land surface properties (e.g. agricultural use) and configuration of the Earth's orbit. The first two of these differences, long timescale and changing boundary conditions, rule out the use of the historical record for model calibration, which shifts the focus from statistical to physically based models so as to explore the effect of boundary conditions different to today, while the third adds a further 'dimension' of uncertainty over which risk assessments must be made. These challenges are described in more detail in the next section.

6.2.1 Challenges in predictability

Prediction is 'very difficult, especially if it's about the future', as the physicist Niels Bohr is said to have pointed out, and the earth system has many particularly challenging features. Risk assessment for hydrometeorological hazards under future climate change has all the difficulties described in Chapter 5 added to those of predicting the distant future. In fact, it might at first seem faintly ludicrous to assert that statements can be made about the earth system decades from now, when weather forecasting has a predictability timescale of about a fortnight and seasonal forecasting is limited to about a year (Chapter 5). But long-term climate change concerns statistical properties rather than chronological order: it is a forced (boundary condition) problem, a problem of the 'second kind' (Lorenz, 1975), rather than one of initial conditions. Chaos (Chapter 5) is therefore not one of the principal difficulties, at least in long-term prediction. Instead, the focus is on predicting, or alternatively defining a plausible range of, future boundary conditions.

Recently there have been increased efforts to make short-term ('decadal') forecasts of year-to-year changes (such as ENSO: Chapter 5) over the next few years. These are timescales of great interest for decision-making and policy-making. But decadal predictions are extremely challenging, because the timescales are too short to be dominated by boundary conditions, and too long to be dominated by initial conditions (Hawkins and Sutton, 2009).

Like seasonal forecasts (Chapter 5), they rely on accurate initialisation to make use of persistent and oscillatory modes of the climate. This newly emerging research field is showing much promise (Smith *et al*., 2007). An overview is given by Hawkins (2011).

Challenges in predicting future boundary conditions (forcings) are described in Section 6.2.1.1; in the earth system response to these forcings in Section 6.2.1.2; and some hazard- and impacts-specific challenges in Section 6.2.1.3.

6.2.1.1 Forcings

A climate forcing is something that acts to perturb the radiative equilibrium of the atmosphere (with units of watts per square metre): by altering either the amount of incoming solar radiation (insolation), or the fraction of radiation that is reflected, or the amount of longwave radiation emitted to space (IPCC, 2007a). These perturbations cause warming or cooling of the atmosphere and/or surface, which is then propagated throughout the earth system. The term 'forcing' is often used interchangeably with '(change in) boundary conditions' because the latter induces the former.

Natural forcings are always acting on the earth system. The amount and distribution of solar radiation reaching the earth varies continuously, in periodic cycles and longer-term changes. Volcanic eruptions occur randomly, emitting sulphur dioxide gas which combines with water in the atmosphere to make sulphate aerosol particles that reflect solar radiation and change the reflectivity (albedo) and lifetime of clouds. Some of these natural forcings are predictable – in particular, the effect of the earth's orbital cycles and the length (though not amplitude) of the 11-year sunspot cycle – but others are unpredictable, such as long-term solar variations and volcanic eruptions.

Anthropogenic (human-caused) forcings arise from large-scale industrial, agricultural and domestic changes to the earth's atmosphere and land surface. Some important forcings are: greenhouses gases (GHGs: including carbon dioxide, methane, nitrous oxide and ozone), which decrease the amount of longwave radiation emitted by the earth to space and have a warming effect at the surface; industrial sulphate aerosol particles, which act similarly to those from volcanoes; and replacement of forests with crops, which increases the reflectivity of the surface and increases the GHG concentrations in the atmosphere, notably if the deforestation involves burning.

Future changes in anthropogenic forcings depend on population, policy, economics and growth, technology and efficiency, mitigation (Section 6.3.1.1) and geoengineering (Section 6.3.1.2). Geoengineering is intentional alteration of the earth system, particularly to counter anthropogenic climate change, with measures that reduce positive (warming) forcings, such as GHG extraction from the atmosphere, or increase negative (cooling) forcings, such as addition of reflective aerosol particles to the atmosphere (Irvine and Ridgwell, 2009; Lenton and Vaughan, 2009). In the short-term, GHGs and industrial aerosol emissions are somewhat predictable, because significant changes in infrastructure, behaviour and population growth take years or decades, and some GHGs have mean lifetimes of several decades. In the longer term, strong mitigation or industrial growth could occur and this is virtually impossible to predict at the centennial scale.

Limited predictability of future anthropogenic forcings is addressed by making predictions for a set of 'possible futures' (emissions scenarios: Section 6.2.4.1). These predictions, which are conditional on a given scenario ('what would happen if'), rather than attempting to encompass every aspect of the future ('what will happen'), are better described as 'projections', though the distinction is not always made. Forcing scenarios usually attempt to encompass the range of plausible future human influence on climate, because the motivation for their use is assessment of risk management strategies. Unpredictable natural forcings are usually set at present-day levels (solar output) or not included (volcanic eruptions).

6.2.1.2 Response to forcings

The statistical properties of weather are challenging to simulate in the present day (Chapter 5), let alone the future. Not only are natural and anthropogenic forcings always changing, but some parts of the earth system take centuries (the deep oceans), millennia (the cryosphere) and longer to respond so the earth system is never at equilibrium: in other words, stationarity can never be assumed except on short temporal and spatial scales. This means the earth's future response to forcings cannot be deduced or extrapolated from the past. Even if forcings were to stay fixed at present-day levels, an assumption of stationarity would have limited applicability.

Feedbacks (Chapter 5) modulate the climate response to forcings. The distinction between forcing and feedback may depend on the timescale of the processes under consideration: for example, on short timescales land vegetation and ice sheets might be treated as (approximately) static boundary conditions, but on long timescales as dynamically responding components of the earth system. Cloud–climate feedbacks are the main source of uncertainty in the time dependent response of surface temperature in climate models (Soden and Held, 2006), particularly in the shortwave radiation feedback of low-level clouds (Bony and Dufresne, 2005; Crucifix, 2006; Webb *et al.*, 2006). Feedbacks between vegetation, climate and carbon dioxide (CO_2) affect surface water runoff, soil moisture and carbon storage, but the magnitude and sign of the response vary between models (Cox *et al.*, 2000; Betts *et al.*, 2007; Alo and Wang, 2008). Some known positive feedbacks, such as methane clathrate release (Maslin *et al.*, 2010), are not routinely, or ever, included in predictions, which may lead to underestimation of future change (Scheffer *et al.*, 2006; Torn and Harte, 2006).

Modelling of climate extremes and their impacts is either combined (known as 'coupled', 'interactive' or 'dynamic') or separate ('offline', 'prescribed'), so the distinction between them can be somewhat subjective. The arbitrariness of these boundaries is largely due to the presence of feedbacks: for example, the two-way interactions between climate and vegetation are now much more commonly included in climate models, though in other studies the impacts of climate and weather extremes on vegetation are modelled offline (i.e. the feedback loop is cut). Feedbacks between hazards and impacts can either amplify or diminish the latter. Dlugolecki (2009) gives examples for insured losses from flooding: these may be diminished (negative feedback) with risk management by reduced habitation of at-risk areas or improved building structures, or amplified (positive feedback) in large-scale disasters by

long-term neglect of properties (after evacuation) or political pressure on insurers to pay for uninsured losses and property improvements. Human-related feedbacks, such as the interactions between climate and the global economy, are represented in integrated assessment models (e.g. Leimbach *et al*., 2010), though the climate modules are usually much simpler than the state-of-the-art.

6.2.1.3 Hazard- and impact-specific challenges

All weather hazard events are challenging to predict because of their rarity (limiting the number of observations and the ability to understand the relevant processes) and their position in the extremes of chaotic internal variability (Chapter 5). Some types are particularly hard to predict. For future changes, extreme temperature is thought the most straightforward, because climate models simulate the major driving processes (Chapter 5; IPCC, 2012), though the regional effects of feedbacks and changes in atmospheric and ocean circulation on temperature extremes are still poorly understood (Clark *et al*., 2010). Extreme precipitation and tropical cyclones are far more challenging, because the processes and weather systems occur on small spatial scales (Knutson *et al*., 2010; IPCC, 2011): for example, while warm sea surface temperatures drive the genesis of tropical cyclones, other factors such as vorticity and wind shear are thought to be important in determining their frequency (Kim *et al*., 2010).

Predictions for some hazards do not even agree on the sign of change in many regions of the world (IPCC, 2012), in particular for multivariate hazards (such as drought) and modes of variability (such as ENSO). One of the difficulties with multivariate hazards is in their definition (IPCC, 2012): Burke and Brown (2008), for example, find the magnitude and sign of predicted drought changes depend on the drought index (Chapter 5) used. Even if predictions of the main drivers of drought (temperature, precipitation, evaporation and runoff) agree, predictions of the resulting soil moisture may be wildly different (Dai, 2011).

6.2.2 Physical modelling

The previous chapter described how physical models or 'simulators' of the earth system are used to simulate weather. Ensemble forecasts, generated by perturbing initial conditions and other uncertain quantities, are used to approximate probability distributions for hazard events. Ensemble forecasts are repeatedly tested against observations to assess whether the ensemble is 'well-calibrated' (explores all important uncertainties). Statistical modelling of past observations is also used, for seasonal forecasting and for estimating return periods, but this is not possible for the longer term: changing boundary conditions place physical modelling at the heart of predicting changes in extremes.

For long-term risk assessment of extremes, then, physical models are run for future scenarios (to make predictions) and for the past (to critique the model). This section describes these models and the difficulties with using them to make multi-decadal simulations of climate: their complexity, computational expense and inevitable limitations in

representing the earth system. Added to these is the multi-decadal timescale of climate, meaning predictions cannot be repeatedly tested against observations to ensure a model is adequate and an ensemble well-calibrated, and the unpredictable future boundary conditions (Sections 6.2, 6.2.1). There are several approaches to overcoming these difficulties, involving hierarchies of physical models (described in this section) supplemented with statistical modelling (described in the next), along with expert judgement and other strategies. However, many of these methods are still under active development, and a consensus on approaches is far from emerging.

The most advanced, general circulation models (GCMs), are similar to numerical weather prediction (NWP) models, but have lower spatial and temporal resolution and include different processes (more details are given in Chapter 5). A small number of models, including the UK Met Office Unified Model, are used for both climate prediction and NWP. Climate models are used for much longer simulations (decades to centuries) than NWP models, so they incorporate more long-timescale processes such as deep ocean circulation, atmospheric chemistry and the carbon cycle. Components with very long time-scale responses, in particular the cryosphere, can also be included, though the coupling is usually asynchronous to reduce computational expense: for example, one-year simulations from the atmosphere–ocean model are alternated with ten-year simulations from the ice-sheet model. The most comprehensive GCMs, particularly those that include chemical and biological as well as physical processes, are also called 'Earth System Models'. The simulation length and complexity of GCMs result in significant computational expense. Lower resolution helps to mitigate this, though simulation rate may still be as slow as 50 model years per month, even on a supercomputer (Easterbook and Johns, 2009). A reference text on climate modelling is McGuffie and Henderson-Sellers (2005); an overview of their current abilities is given by the US Climate Change Science Program (CCSP, 2008); and the current state-of-the-art, multiple climate model dataset (CMIP5: Section 6.2.4.3) is described by Taylor *et al*. (2011).

Recently there has been greater emphasis on making predictions with a continuum of resolutions and timescales, from three-day weather predictions to decadal and centennial climate predictions, in order to share their strengths: for example, the climate prediction configuration of the Unified Model benefits from daily testing of the operational weather forecasting configuration. This is termed 'seamless prediction' (Rodwell and Palmer, 2007; Palmer *et al*., 2008).

The low resolution of GCMs, hundreds of kilometres horizontally (Chapter 5), places great reliance on the parameterisation of small-scale processes (Section 6.2.2.1). A spectrum of model complexity exists, in which processes are represented in parameterised form to a greater or lesser extent (Section 6.2.2.2). Simpler, faster models are used for longer simulations and large ensembles, and complex models for shorter simulations, small ensembles or reduced domains. Predictions of regions (country to continental scale) suitable for impacts studies are made with high-resolution, limited-area models similar to those of operational NWP (regional climate models: RCMs), driven by boundary conditions derived from climate models. This is known as physical downscaling (Section 6.2.2.3).

6.2.2.1 *Parameterisation*

Physical models can never include every process, describe all included processes perfectly or have infinite resolution, so the behaviour of imperfectly calculated processes is represented using parameters equivalent to real-world or abstracted quantities. These uncertain parameters are given fixed standard values determined from calibration with observations, expert knowledge and specific high-resolution physical modelling studies (e.g. Randall *et al.*, 2003). An important example is convective cloud formation, where the relevant processes occur on sub-grid spatial scales. The clouds must be represented with bulk formulae: for example, relating fractional cloud cover in a grid box to relative humidity. Other examples involve interactions at the earth's surface, such as energy and moisture fluxes between the boundary layer, land and sea ice. Important parameterisation schemes in climate models are reviewed by CCSP (2008) and McFarlane (2011). A reference text is Stensrud (2009).

Model calibration (tuning) aims to find parameter values that bring the model simulation in closest agreement with the target datasets for a wide range of physical quantities. This is usually an impossible aspiration because the most successful values of some parameters are incompatible with those of others, or because improvements in simulating one physical quantity come at the expense of others. Climate models are tuned to reproduce mean observations, not extremes, due to the limited number of observations. Another important requirement of tuning (not necessary for NWP models) is to keep the global energy balance at the top of the atmosphere stable over long periods of time, to avoid artificial temperature drift (CCSP, 2008). Parameter uncertainty in climate simulations is explored by 'detuning' the model. The resulting ensembles are used directly or re-weighted with comparisons to observations (Section 6.2.4.2).

An alternative to parameter calibration is the ASK method (Stott and Forest, 2007), rooted in detection and attribution analysis, in which a model prediction is re-scaled by the error compared with historical observations; the model response to each type of forcing is separated out and scaled separately. The ASK method uses an assumption of a linear relationship between errors in simulating past climate change and predicting future climate change, and is most reliable if the relative fractions of different forcings remains the same (Stott and Forest, 2007); this is not the case for mitigation or geoengineering scenarios (Section 6.3.1). The ASK approach has been applied to regional change (Stott and Forest, 2007), but attribution methods for extremes are in their infancy (e.g. IPCC, 2012; Min *et al.*, 2011; Pall *et al.*, 2011).

Most climate models are deterministic, with fixed parameter values. But physical consistency can fail for deterministic parameterisations: for example, violating conservation of energy at small spatial scales. It is increasingly recognised that stochastic components, introduced through randomly varying parameters or processes (Chapter 5) can improve climate model predictions (Palmer, 2001; Palmer *et al.*, 2005; Palmer and Williams, 2009; McFarlane, 2011).

6.2.2.2 *Spectrum of model complexity*

Physical climate models vary from complex, computationally expensive GCMs and RCMs, to models of intermediate complexity that have even lower resolution and a larger fraction of

parameterised processes and prescribed components (earth system models of intermediate complexity: EMICs), to simple conceptual ('toy') models. Some of this variation is a reflection of continuing model development over time, with state-of-the-art model versions used alongside older, faster versions, and some is due to design, with simpler models used to improve understanding or for low computational expense. A model is selected and configured according to the scientific problem and available resources: long-timescale processes may be neglected if simulations are short; simpler, faster models enable many replications for statistical analysis. In general, simple and intermediate complexity models cannot adequately simulate regional climate change and extremes (Stott and Forest, 2007), but they may be tuned to GCM or RCM simulations to scale their results for other forcing scenarios (Section 6.2.4.1). Even state-of-the-art models do not simulate some extremes well (Section 6.2.1.3): precipitation is one example for which models typically underestimate the frequency and intensity of extreme events (IPCC, 2007a).

Physical models of the impacts of extreme weather (health, environment and financial: Chapter 5) also exist for some research areas. Environmental models of varying complexity exist for the effects of drought, extreme temperature and extreme precipitation on crops (e.g. Challinor *et al*., 2004), forests (Allen *et al*., 2010), terrestrial ecosystems (e.g. Knapp *et al*., 2008) and interactive vegetation–climate feedbacks (e.g. JULES: Best *et al*., 2011; Clark *et al*., 2011). Vegetation models can be adapted to simulate the impacts of heat waves on cities (Masson, 2006), and bio-physical models simulate the impacts of heat waves on health with human energy balance equations (McGregor, 2011). Physical models of structures also exist: for example, a model of the impacts of extreme precipitation on urban drainage systems is described by Wright *et al*. (2006). The significant advantage of physically based impacts models is that they do not assume the relationship between hazard and impact is fixed in the future: for example, the effect of changing carbon dioxide concentrations on crops is incorporated in vegetation models. However, physically based models do not exist for all types of impacts and this is a rapidly progressing field.

6.2.2.3 Dynamical downscaling

'Downscaling' means increasing the spatial or temporal resolution of a dataset, usually the output of a GCM. The term 'spatial downscaling' is generally reserved for methods that incorporate the effects of local topography, coastlines and land use on the signal, rather than simple linear interpolation to a higher-resolution grid. Downscaling is crucial for hazard impact assessment, because the hazards and impacts occur on much smaller spatio-temporal scales than GCM resolutions (e.g. extreme temperature and precipitation: Diffenbaugh *et al*., 2005; Gosling *et al*., 2009; crops: Tsvetsinskaya *et al*., 2003). There are two very different approaches, based on statistical (Sections 6.2.3.2, 6.2.3.3) and physical ('dynamical') modelling.

Dynamical downscaling is another name for regional climate modelling (proprietary versions have also been developed for catastrophe loss modelling: Section 6.2.3.2). RCMs are very computationally intensive, which restricts the length and number of simulations, but unlike statistical downscaling they generate physically and spatially consistent

results for a large number of variables and allow changes in boundary conditions to be made. For predictions of future climate, an RCM is either run offline with boundary conditions (such as temperature, wind and pressure) at the domain edges (and, if the RCM is used to simulate only the atmosphere, the ocean surface) from a pre-existing GCM simulation, which only allows movement of weather systems from the GCM to the RCM ('one-way nesting'), or coupled to a GCM, allowing weather systems to move between the domains ('two-way nesting'). The former is predominant, being more straightforward technically and computationally, and useful for comparing RCMs, but the latter is more physically realistic (e.g. Lorenz and Jacob, 2005). Variable-resolution global models are also used (Fox-Rabinovitz *et al.*, 2008). Downscaling is reviewed by the CCSP (2008) and Rummukainen (2010).

Tropical cyclones are not resolved by the current generation of RCMs. They are simulated with tailored high-resolution physical models or statistical models driven by climate model simulations (Chapter 5).

6.2.2.4 Extremes and anomaly methods in physical modelling

GCMs are a 'necessary evil' in estimating the effect of changing boundary conditions on extremes, because they incorporate causal physical mechanisms that relate global forcing (e.g. CO_2 emissions) to regional climate (e.g. the distribution of European daily temperatures). The response of climate to forcing is estimated either with two steady-state (equilibrium) simulations – for example, one using pre-industrial boundary conditions and the other using doubled atmospheric CO_2 concentrations – or with a time-evolving (transient) simulation of the historical period and next one or two centuries. Climate simulations from GCMs are downscaled with regional climate modelling (previous section), weather generation (Section 6.2.3.2) or statistical spatial downscaling (Section 6.2.3.3) to increase their temporal and spatial resolution. This is necessary for estimating extremes at a given location and for providing inputs to impacts models.

The frequency and intensity of extremes in downscaled simulations can be quantified empirically (e.g. counting threshold exceedances) or, particularly for rarer extremes, modelled statistically (e.g. fitting a GPD model to those threshold exceedances: Section 6.2.3.1). These methods either use an assumption of stationarity (for equilibrium simulations or short sections of transient simulations) or quantify the time-evolving change (e.g. by allowing the parameters of the statistical model to vary with time).

Climate simulators inevitably have systematic biases. When studying the effect of changing boundary conditions on extremes, some of these may be cancelled out, to first order, by focusing on the change ('anomaly') between the simulations of the present-day and future scenarios, rather than the absolute values of the latter. Effectively the future distribution of extremes is inferred from a combination of the present-day observational record, present-day simulation and future scenario simulation (Ho *et al.*, 2012). There are two methods, which are exactly equivalent in a special case: if the distributions of extremes in the three datasets all have the same scale and shape.

The first is 'bias correction', in which the future scenario simulation is corrected by the model discrepancy, which is the difference between the present-day simulation and

observations. The second is the 'change factor' method, in which observations are changed by the simulated climate anomaly, which is the difference between the present-day and future simulations. The simplest approach to these is to apply only the mean change (mean discrepancy or mean anomaly). This is appropriate if the distributions in the three simulations have the same scale (e.g. variance) and shape (e.g. skewness). But typically they do not, so a better approach is to apply the changes to several quantiles (for empirically estimated distributions) or to all the parameters (for parametric approaches such as GPD modelling or statistical downscaling). Unfortunately, even if this is done, bias correction and change factors can give very different results (Ho *et al.*, 2012). The appropriate choice, including the underlying assumptions, is not obvious and this is an important area of future research.

6.2.3 Statistical modelling

Statistical modelling has two supporting roles in estimating future extremes from physical climate models such as GCMs. The first is to characterise properties of the extremes in a given simulation: to estimate return periods (extreme value analysis, Section 6.2.3.1) or generate pseudo-event datasets for impacts studies (weather generation and cat modelling, Section 6.2.3.2). The second is to supplement and extend a set of simulations, by quantifying relationships between high- and low-resolution information (statistical spatial downscaling, Section 6.2.3.3) or model parameters and output (emulation, Section 6.2.3.4).

Many impacts models are statistically based, because the mechanisms are not known (Chapter 5). These implicitly assume the relationship between hazard and impact (e.g. mortality) is fixed in the future, because the changing effect of external factors such as economic growth and infrastructure is impossible to predict. Catastrophe ('cat') models of loss, developed for insurance and reinsurance, incorporate some physical-based constraints but are principally statistical.

6.2.3.1 Extreme value analysis

Methods of analysing extreme values in observational records are described in the previous chapter: for example, fitting a generalised extreme value (GEV) distribution to annual maxima, or a generalised Pareto distribution (GPD) to threshold exceedances. These methods estimate long-term return periods for extremes of a given magnitude, usually extrapolating to rarer extremes than are in the observations and using an assumption of stationarity.

These extreme value theory (EVT) methods are increasingly used for analysing climate simulations (Naveau *et al.*, 2005; Katz, 2010; references therein). EVT relies on an asymptotic limit (Chapter 5): in principle, GCMs could be used to generate very long equilibrium simulations, tens of thousands of years, to approximately satisfy this assumption, but in practice the models are too computationally expensive to allow simulations of more than 10–20 decades (the problem is worse for the observational record: Chapter 5).

Extremes are either assessed in two equilibrium simulations, two near-equilibrium parts of a simulation or continuously through a transient simulation. An example of the second is to estimate the change in GEV location parameter (Chapter 5) in simulations of the present day and the end of the twenty-first century. Bias correction or change factor methods (Section 6.2.2.4) are used to remove some systematic simulator errors, and hypothesis tests (e.g. goodness-of-fit tests: Chapter 5) are applied to determine the significance of any change. Examples of the third, analysing extremes in non-stationary data, are described in Chapter 5: for example, GEV or GPD parameters may be allowed to vary through the simulation in step-changes (e.g. Min *et al*., 2009), with a linear trend (e.g. Cooley 2009; Bjarnadottir *et al*., 2011) or covarying with global mean temperature or an index of natural variability (e.g. Hanel *et al*., 2009; Hanel and Buisand 2011).

6.2.3.2 Weather generation and catastrophe modelling

Weather generation is statistical modelling of a dataset to create many plausible realisations of the same time period with the same statistical properties as the original (e.g. Qian *et al*., 2004; Kyselý and Dubrovský, 2005; Semenov, 2008; Jones *et al*., 2009; Kyselý, 2009; Nicolosi *et al*., 2009). This is a type of statistical downscaling applied to GCM or RCM output to increase temporal resolution. In parametric weather generation, the parameters of a time-series model are calibrated with observations or climate simulations; non-parametric weather generators use bootstrapping methods (Chapter 5). Each location is usually treated independently (Chapter 5). The pseudo-hazard datasets are used for risk assessments or as inputs to impacts models; they may need further processing for this (e.g. thermal simulation of buildings: Eames *et al*., 2011). Bias correction or change factor methods (Section 6.2.2.4) are used to make projections: in the latter, parameters of the weather generator are calibrated with observations then modified by projected climate anomalies from the GCM (e.g. Jones *et al*., 2009). For changes in extreme weather it is important to apply not only mean climate anomalies but also aspects of variability such as variance, skewness and correlation. An overview of weather generation is given by Wilks and Wilby (1999), and a more recent discussion in the context of extreme precipitation is given by Furrer and Katz (2008).

Catastrophe models (or 'cat models') of hurricanes and other hazards are statistical methods for generating pseudo-event datasets that incorporate the whole chain from hazard to financial loss (Chapter 5). In the past, cat models have been based on sampling or calibration with historical observations, implicitly assuming stationarity, but recently RCMs (Schwierz *et al*., 2010) and proprietary numerical weather models (the 'US and Canada Winterstorm' and 'European Windstorm' models developed by Risk Management Solutions) have been incorporated into loss modelling so as to make risk assessments under changing boundary conditions.

6.2.3.3 Statistical spatial downscaling

Statistical spatial downscaling, sometimes referred to as statistical bias correction, uses empirical relationships between information at high spatial resolution (meteorological

station data, such as precipitation) and low spatial resolution (GCM or mean station data, such as temperature, humidity and sea-level pressure). These relationships are derived using statistical modelling techniques such as regression (e.g. Wilby *et al.*, 2002), artificial neural networks (e.g. Mendes and Marengo, 2009) and synoptic classification schemes. Statistical temporal downscaling is known as weather generation (Section 6.2.3.2). Techniques for downscaling mean climate may not be suitable for extremes: for example, extremal indices (Chapter 5) are often non-Gaussian, so the underlying assumptions of linear regression are not valid and a more flexible form such as a generalised linear model may be more appropriate. Benestad *et al.* (2008) describe strategies and example code for downscaling extremes and probability distributions.

These methods are easy to implement and, compared with dynamical downscaling (Section 6.2.2.3), do not require the very significant computational expense, can access finer scales, are applicable to variables that cannot be directly obtained from RCM outputs and can be comparably successful (IPCC, 2007a; CCSP, 2008). However, they cannot include feedbacks, there is no spatial coherence between sites, physical consistency between climate variables is not ensured and, most importantly, they include an implicit assumption that these relationships remain unchanged under different boundary conditions (IPCC, 2007a).

If meteorological data with good spatial coverage are available, an alternative is to use a change factor approach (Section 6.2.2.4), adding GCM climate anomalies to an observed climatology (e.g. Hewitson, 2003). Using a mean change factor does not allow for changes in variability and temporal correlation and makes the assumption, like pattern scaling (Section 6.2.4.1), of an unchanging spatial pattern; for non-Gaussian variables such as precipitation, care must be taken to avoid unphysical results (Wilby *et al.*, 2004).

6.2.3.4 Emulation

Statistical representation of complex models, known as 'emulation' (or 'meta-modelling'), is used to make predictions for areas of parameter space not explored by a physical model (O'Hagan, 2006). Emulators are usually constructed using regression, with coefficients calibrated from an ensemble of simulations with different parameter settings (e.g. Murphy *et al.*, 2004, 2007; Kennedy *et al.*, 2006; Rougier and Sexton, 2007; Rougier *et al.*, 2009; Sexton *et al.*, 2011), though neural networks have also been used (e.g. Knutti *et al.*, 2003; Piani *et al.*, 2005; Sanderson *et al.*, 2008). The primary motivation is estimation of simulator uncertainty for models that are too computationally expensive to allow many replications. Expert judgement may also be incorporated by requiring that the outputs satisfy physical principles such as conservation of mass and water. Emulation is conceptually simple, but requires careful choices to avoid problems such as over-fitting (Rougier *et al.*, 2009; Sexton *et al.*, 2011). Success of an emulator can be assessed with diagnostic tests, such as leaving one or more model simulations out of the emulator calibration and predicting their output (Rougier *et al.*, 2009; Sexton *et al.*, 2011).

6.2.4 Uncertainty assessment

The previous two sections describe the difficulties with simulating long-term changes in extreme weather. Computational limitations restrict the spatial and temporal resolution, so that their realism and relevance is compromised, and this can only partially be addressed by parameterisation and tuning. Our understanding is incomplete for the physics, chemistry and biology of some aspects of the earth system, such as the cryosphere or vegetation, that might well be important for 100-year time intervals under changing boundary conditions. So, although climate simulators can be run in ensembles to approximate probabilities of extreme weather events, these approximations are very crude, being based on a small number of runs, and compromised by the limitations of the code. On top of this, future boundary conditions are unknown, adding a further uncertainty to be sampled.

The physical and statistical modelling methods described can address some of these issues of model inadequacy: bias correction or change factors to cancel some systematic errors; a hierarchy of physical model complexities and statistical emulation for different research questions (high complexity for improving representation of extremes; low complexity and emulation for increasing simulation length or number); physical and statistical downscaling to increase resolution; and extreme value analysis to extrapolate to rarer extremes than can reasonably be simulated. This section further describes these methods, and other tools, on the topic of quantifying uncertainties, including: ensembles that sample uncertainties in initial and boundary conditions, model parameters and structure; pattern scaling and emulation to extend ensemble size; introduction of a 'discrepancy' term to account for structural uncertainty; and approaches to calibration of parameters and discrepancy using the historical record.

The future of the earth system cannot be known exactly, because future forcings are unpredictable, but it is of paramount importance to assess uncertainties of climate model projections conditional on these forcings to inform decision-making and policy-making on potential options for mitigation (Section 6.3.1.1) and adaptation (Section 6.3.2). Uncertainties for extreme weather hazard projections can be partitioned into two kinds, associated with physical model predictions and the statistical modelling of those predictions. The first kind arise from uncertain physical model inputs (Section 6.2.4.1), parameters (Section 6.2.4.2) and structure (Section 6.2.4.3), and are explored and quantified using ensembles of predictions using different inputs (boundary and initial conditions), parameter values and physical models (an overview is given by Foley, 2010). The second kind are from statistical assessment of simulations and from statistical prediction of physical model output such as emulation. Uncertainties in statistical model parameters and structure are assessed with the usual methods, including graphical checks, bootstrapping and hypothesis tests, described in Chapter 5.

6.2.4.1 Model input uncertainty

Boundary condition uncertainty is sampled with a suite of predictions for different future 'storylines' (Section 6.2.1.1), scenarios that describe the effect on anthropogenic emissions

of plausible future societal and technological changes, including economic growth, energy sources and efficiency and land use. Forcing scenarios are not predictions, and do not have relative probabilities assigned, so climate predictions are made conditionally upon each individual scenario. For the IPCC Fourth Assessment Report (IPCC, 2007a), projections are made based on the Special Report on Emissions Scenarios (SRES: Nakićenović *et al.*, 2000), which included storylines that ranged from rapid growth and fossil-intensive energy sources ('A1B') to global sustainability approaches ('B1'); the storylines are detailed descriptions of GHGs and sulphur dioxide emissions, which are translated into atmospheric concentrations using biogeochemical modelling (IPCC, 2007a). For the Fifth Asssessment Report the focus is 'representative concentration pathways' (RCP: Vuuren *et al.*, 2011); unlike SRES, these include scenarios with explicit climate policy intervention (Section 6.3.1.1) and are simpler, being expressed in terms of atmospheric concentrations so they can be used directly in climate models. Forcings due to possible geoengineering strategies (Section 6.3.1.2), including changes to stratospheric albedo (e.g. Ricke *et al.*, 2010), land surface albedo (e.g. Ridgwell *et al.*, 2009; Oleson *et al.*, 2010) and carbon uptake (e.g. Aumont and Bopp, 2006), are considered in separate studies. Future natural forcings (solar, volcanic) are 'known unknowns'. An early suite of climate projections (Hansen *et al.*, 1988) incorporated volcanic eruptions, but these are not usually included and no attempt is made to assess the effect of changes to solar forcing from the present day.

Climate models are so computationally expensive that simpler models are often used to approximate their predictions to increase the size of the boundary condition ensemble. This is known as 'pattern scaling': a simple climate model (SCM) simulates global mean temperature as a function of atmospheric forcing (i.e. for a range of scenarios), and GCM results from one scenario are scaled by the global mean temperature from the SCM. Pattern scaling relies on the assumption that simulated patterns of climate change scale linearly with global temperature; this is thought to be less reliable for precipitation than for temperature (Mitchell, 2003; IPCC, 2007a; Wilby *et al.*, 2009). Harris *et al.* (2006) estimate pattern scaling error by comparing the result of a 'slab' (simplified ocean) GCM scaled by a simple model with 'truth' from the dynamic ocean version of the GCM.

Initial condition uncertainty – incomplete knowledge of the initial state of the system – is also sampled with ensembles of simulations (Chapter 5). It is most important in decadal climate predictions, in which internal variability dominates (e.g. Smith *et al.*, 2007). Long-term projections focus on trends rather than year-to-year changes, so initial condition ensembles for these tend to be small and averaged over (e.g. Stainforth *et al.*, 2005; Meehl *et al.*, 2007) or else not performed at all.

6.2.4.2 Model parameter uncertainty

Uncertainty in climate model parameterisation schemes can be explored using perturbed parameter ensembles (PPEs), more usually and misleadingly named perturbed *physics* ensembles, which are groups of 'detuned' model versions created by changing the values of the control parameters (Section 6.2.2.1). Perturbed parameter ensembles offer the opportunity to sample uncertainty systematically in a well-defined space. They are often

supplemented with statistical methods such as emulation (Section 6.2.3.4) to improve this sampling. A further motivation is to test detuned model versions against past observations, to avoid the circularity induced by tuning and validating with the same time period (Section 6.2.4.3). The full, unweighted ensemble of simulations may be analysed, or if the parameters are sampled from probability distributions (e.g. uniform, Gaussian) with sufficient sample size they can be integrated out using Monte Carlo methods to assess the contribution of parametric uncertainty to the predictions. The latter method can include likelihood weighting of ensemble members in a Bayesian statistical framework (Chapter 5) by comparing simulations of the recent past with observations (e.g. Sexton *et al.*, 2011).

The two most extensively studied PPEs use variants of the Hadley Centre climate model, which is a version of the UK Met Office Unified Model. The first is an ensemble of 17 versions, supplemented by an ensemble of about 300 using a simplified ocean model, used to generate projections for the UK Climate Impacts Programme (UKCIP: Murphy *et al.*, 2009; Sexton *et al.*, 2011). The second is a suite of ensembles with different model complexities, domains and experimental aims, each with hundreds to tens of thousands of members, generated with distributed (publicly donated) computing for the project *ClimatePrediction.net* (Stainforth *et al.*, 2005 and many subsequent studies). PPEs have also been performed with other models than the Hadley Centre model (e.g. Yang and Arritt, 2002; Lynn *et al.*, 2008; Sterl *et al.*, 2008; Yokohata *et al.*, 2010; Fischer *et al.*, 2010; Klocke *et al.*, 2011; Sanderson, 2011). PPEs can be performed with RCMs, but their computational expense limits them to small ensemble size (Murphy *et al.*, 2009; Buontempo *et al.*, 2009), short simulation length (Yang and Arritt, 2002) or both (Lynn *et al.*, 2008). PPEs are also generated with simpler, lower-resolution models, such as EMICs, but these typically focus on global mean temperature rather than regional patterns or extremes. An overview of PPEs is given by Murphy *et al.* (2011).

The original motivation for PPEs was quantification of uncertainty in the global mean temperature response to CO_2 forcing (Andronova and Schlesinger, 2001; Murphy *et al.*, 2004; Stainforth *et al.*, 2005), but they have since been used in risk assessments of future extreme temperature and precipitation (Barnett *et al.*, 2006; Clark *et al.*, 2006; Sterl *et al.*, 2008; Fischer *et al.*, 2010; Fowler *et al.*, 2010), drought (Burke and Brown, 2008; Burke *et al.*, 2010; Hemming *et al.*, 2010) and storms (Kim *et al.*, 2010). The disadvantage of studying extremes in PPEs is that for a given computational resource have to be made with simulation length: for example, the UKCIP simulations using a simplified ocean model are only 20 years long (e.g. Barnett *et al.*, 2006).

Climate model PPEs have not yet been much used in assessments of future hazard impacts. PPE methods have been used to some extent as inputs to impacts models (e.g. heat wave mortality: Gosling *et al.*, 2011), but propagation of the cascading parametric uncertainty in both climate and impacts models – the simplest method being a factorial experiment, with every combination of parameter values – is still relatively rare (e.g. crops: Challinor *et al.*, 2005, 2009).

6.2.4.3 Model structural uncertainty

Structural uncertainty arises from imperfect implementation or knowledge about the relevant physical processes: it may be considered the leftover error, or 'discrepancy', at the model's best (tuned) parameter values (Rougier, 2007). It is represented with a covariance matrix that describes the variations and correlations of model discrepancy across locations, times and variables. There are various methods of estimating this covariance matrix in numerical weather prediction (Chapter 5), which rely on repeated tests of the model forecasts against reality. These methods cannot be directly translated to climate model predictions, because of the multi-decadal definition of climate. Structural uncertainty is therefore extremely challenging to quantify, but necessary for assessing the success of climate model predictions and for decision-making under future climate change. A related problem is the appropriate interpretation of a range of predictions from different climate models. In the past, model structural uncertainty was considered so difficult to quantify it was essentially ignored, and all models treated as equally skilful, but the literature has flourished since the IPCC Fourth Assessment Report (IPCC, 2007a). The main points are summarised here.

Expert elicitation has been proposed as a method of estimating structural uncertainty (Kennedy and O'Hagan, 2001; Goldstein and Rougier, 2004). The state space of climate models is extremely high-dimensional, so many simplifying assumptions about the correlation structure (spatial, temporal and multivariate) would be necessary in order to elicit a covariance matrix and the complexity of GCMs may be too great for this to be practicable (Sexton *et al*., 2011). Instead, structural uncertainty is inferred by comparing model output with expectations, observations and output from other models.

Simulator discrepancy arises from two sources: imperfect implementation of theoretical knowledge (e.g. numerical approximations, coding errors), and imperfect theoretical knowledge of the system (missing or imperfectly represented physical processes). The former are inaccuracies, checked with 'verification', while the latter are inadequacies, assessed with 'validation' (e.g. Roy and Oberkampf, 2011). Verification of climate models is challenging due to their complexity, and true verification may be impossible (Oreskes *et al*., 1994). Both formal and informal methods are used, including bit-wise comparisons with previous versions, isolated testing of numerical methods and frequent comparison with observations and other models, though recommendations have been made to improve and extend these (Pope and Davies, 2002; Easterbrook and Johns, 2009; Clune and Rood, 2011). The UK Unified Model is used for operational NWP so it is tested very frequently, albeit under different configuration and timescales than for climate prediction.

Validation of climate models is even more challenging: the difficulties of their complexity are compounded by the long timescale of climate and the relative sparsity of observations of the high-dimensional, multi-decadal state. The model is evaluated with observations, where available, or with other models. If the model does not agree with observations, after accounting for observational uncertainty, the remaining discrepancy can inform estimates of the model structural uncertainty; conversely, model evaluation (e.g. of members of a PPE)

requires an estimate of the model structural uncertainty to be meaningful (e.g. Sexton and Murphy, 2011). If a model agrees well with observations, it is not guaranteed to be successful in prediction; reasons are described below. Validation can therefore only be a partial assessment of adequacy (Oreskes *et al.*, 1994; Knutti, 2008a). Methods and interpretation are vigorously discussed and debated in the literature (e.g. Räisänen, 2007; Knutti, 2008a; Reichler and Kim, 2008; Hargreaves, 2010; Fildes and Kourentzes, 2011). The main challenges and strategies are discussed here briefly.

Climate modelling as a research field is only four or five decades old, not much longer than the timescale on which climate is defined, so projections of future climate cannot be repeatedly tested against observations to ensure ensemble predictions are well-calibrated (as in NWP: Chapter 5). Some out-of-sample climate forecasts exist and these have appeared to be quite successful (e.g. Smith *et al.*, 2007; Hargreaves, 2010), though these are single rather than repeated tests. Climate model validation therefore relies on multiple comparisons of 'hindcasts' rather than forecasts. Their skill in reproducing historical observations is held to be an indicator of their likely success at predicting future climate. There are a number of difficulties with this approach.

First, the availability of data. The observational record is 100–150 years long for a very limited number of variables, a few decades for most variables, and limited in spatial coverage; satellite records are 1–2 decades long. This is partially addressed with 'space-for-time' substitution: repeated model testing is performed across different locations instead of, or as well as, different times. One example is the recent application to climate model simulations of the rank histogram, a skill metric used in NWP, by replacing multiple forecast tests in time with multiple hindcast tests in space (Annan and Hargreaves, 2010). However, such tests require careful consideration of the spatial correlation structure of the model discrepancy. Datasets of extremes are, of course, particularly limited; evaluation of climate model extremes focuses on their magnitude and trend (CCSP, 2008).

A second difficulty is circularity: climate models are both tuned and evaluated with observations in the historical period. So if a model reproduces the historical observations, the best that can be said is that it is empirically adequate or consistent (Knutti, 2008a, 2008b). There are various strategies to mitigate this circularity. One is to decrease reliance on parameterisation by adding physical processes and increasing resolution, but this brings relatively slow improvements. A second is to consider agreement with observations as a constraint on the model parameter space (Knutti, 2008b) or, more formally, to detune climate models and incorporate this uncertainty in future predictions (PPEs: Section 6.2.4.2). Circularity could also be avoided by validating with independent observations, but these are not easy to come by: datasets are relatively few (CCSP, 2008), and new ones slow to appear (due to the global coverage and long timescale required), and model developers are generally aware of datasets even if they are not formally used in tuning. One solution is to validate the model with independent information from palaeoclimate reconstructions. These reconstructions are based on 'proxies' such as the response of vegetation (inferred from, for example, fossilised pollen or tree rings) to climate; they are more uncertain than observations, but the signal is often large enough to compensate. These datasets have been

used to test the relative success of different climate models (e.g. Braconnot *et al.*, 2007) and perturbed parameter ensemble members (e.g. Edwards *et al.*, 2007). Model parameters have also been tuned to palaeoclimate information (e.g. Gregoire *et al.*, 2010). However, palaeoclimate reconstructions are usually of annual or seasonal mean climate, not extremes, and are not always provided with uncertainty estimates, which are required for these methods.

A third difficulty is the assumption that a model that is successful at hindcasting will also be successful at forecasting. It is proper to assume that the laws of physics are constant, but successful parameter values may depend on the state of the system. Tuning can also conceal modelling inadequacies such as compensating errors or boundary conditions (Knutti, 2008b). This point can also be addressed with palaeoclimate information: the more out-of-sample tests in which a model is found to be 'empirically adequate', the more likely it is to be adequate for the future state. In fact, there are indications from comparisons with palaeoclimate records that GCMs may be too stable (Valdes, 2011).

Model set-up can be changed to quantify the sensitivity of the results: in particular, the effect of changing resolution or domain. One example is a 'Big Brother/Little Brother' experiment, in which a large-domain RCM simulation is filtered to GCM resolution and used to drive a small domain, to compare with the original RCM result (Rummukainen, 2010). Changes in climate model set-up can also be propagated to impacts models, to test their effect on estimated hazard impacts (e.g. Gosling *et al.*, 2011).

But the primary method of comparing model results and exploring model structural uncertainty is with a multi-model ensemble (MME), a collection of climate models produced by different research and meteorological institutes around the world. These are 'ensembles of opportunity', rather than ensembles of design such as those described above. However, simulations generated by these MMEs are systematically designed and analysed: standardised boundary conditions are used and the simulations made publicly available for the World Climate Research Programme Coupled Model Intercomparison Projects (CMIP3: Meehl *et al.*, 2007; CMIP5: Taylor *et al.*, 2011), so as to allow model intercomparisons (e.g. PRUDENCE: Christensen *et al.*, 2002) and multi-model predictions (IPCC, 2007a). MMEs have been used extensively to assess extreme weather hazards under future climate change, including: extreme temperature (e.g. Meehl and Tebaldi, 2004; Weisheimer and Palmer, 2005; Kharin *et al.*, 2007; Fischer and Schär, 2008; Kyselý, 2009), extreme precipitation (e.g. Palmer and Räisänen, 2002; Meehl *et al.*, 2005; Räisänen, 2005; Tebaldi *et al.*, 2006; Kharin *et al.*, 2007), drought (e.g. Wang, 2005; Burke and Brown, 2008) and hurricanes (Emanuel *et al.*, 2008; references therein).

MMEs can be used to estimate structural uncertainty, so long as some assertion is made about their 'exchangeability': that is, that each is judged equally plausible by experts (Sexton *et al.*, 2011). This approach has been used to estimate the structural uncertainty of the Hadley Centre climate model, with the motivation that models in the MME are likely to incorporate at least some of the processes that are missing in the Hadley Centre model: emulation is used to search parameter space for the closest match to each member of the MME (treating each model in turn as 'truth'), the remaining discrepancy calculated and the structural uncertainty covariance matrix estimated from these discrepancies

across the MME (Sexton *et al.*, 2011; Sexton and Murphy, 2011). The authors acknowledge this may be an underestimate due to common model inadequacies (described below) and test the sensitivity of their results to the method.

MMEs are not systematic samples of a well-defined space, so it is not straightforward to integrate over their future predictions in the same way as PPEs (Section 6.2.4.2). Various methods have been proposed for interpreting their predictions for policy-making (Tebaldi and Knutti, 2007; Knutti, 2010; Knutti *et al.*, 2010). First, with expert judgement: an early estimate of the range of climate sensitivity to a doubling of atmospheric CO_2, 1.5–4.5°C (Charney, 1979), was based on the judgement of a small group of experts about the uncertainty represented by a small number of climate model simulations that spanned the range 2–3.5°C, and this estimate remained unchallenged until recently (IPCC, 2007a). Second, as a 'probabilistic' distribution, analogous to NWP probabilistic forecasts (e.g. Räisänen and Palmer, 2001; Palmer and Räisänen, 2002). Third, as an unweighted mean: this has been found to outperform individual ensemble members (e.g. Lambert and Boer, 2001; Palmer *et al.*, 2004). Fourth, as a weighted mean, based on comparisons with observations (Giorgi and Mearns, 2003; Tebaldi *et al.*, 2005). However, ensemble averaging tends to produce smooth fields biased towards the climate mean state, so it may not be suitable for extremes (Palmer and Räisänen, 2002).

There are very many metrics of skill, and no model in the global MME performs best in all, or even the majority, of these metrics (IPCC, 2007a). But the main difficulty in combining MME predictions is that they are not independent: model components are inevitably similar due to the use of the best available theoretical knowledge, and its representation in code, as well as the movement of model developers and code between institutes. This results in common biases between models (Collins *et al.*, 2011). The degree of independence is extremely difficult to quantify, though several recent attempts have been made (Jun *et al.*, 2008; Pirtle *et al.*, 2010; Collins *et al.*, 2011; Masson and Knutti, 2011; Pennell and Reichner, 2011). A weighted mean based on skill metrics may be worse than an unweighted mean, if inter-model correlations are not taken into account (Weigel *et al.*, 2010).

6.3 Risk Management

Risk management for climate change is an extremely broad topic. Aspects relevant to changing risks of extreme weather are the main focus here; managing the risks of extreme weather hazards in the present climate (disaster risk management) is discussed in Chapter 5. One of the biggest challenges for extreme weather hazards in the medium- to long-term future is the mismatch in timescales under consideration between decision-makers and scientists: from the seasonal to multi-annual time frame of political, insurance and business concerns to the multi-decadal time frame for which climate change can be estimated (Dlugolecki, 2009; IPCC, 2012).

Mitigation and geoengineering are strategies that attempt to reduce the severity of climate change by modifying the projected future forcing (Section 6.3.1). The relative effectiveness of different strategies is estimated with the methods described in this chapter: for example, comparing a 'business-as-usual' emissions scenario with a mitigation scenario that strongly reduces emissions, or comparing a 'constant solar insolation' scenario with one that reduces insolation to approximate the effects of possible future geoengineering. Adaptation strategies attempt to reduce the severity of the impacts of climate change by modifying the future loss (Section 6.3.2). The relative effectiveness of different strategies is estimated with the methods described in this chapter, comparing 'no adaptation' losses with 'adaptation' losses. Communication strategies attempt to inform decision-makers and the public of the projected risks (Section 6.3.3).

6.3.1 Mitigation and geoengineering

The risks of extreme high temperature hazards (heat waves and droughts) increase, of course, with global temperatures. For extreme precipitation and storm hazards, atmospheric and ocean temperatures are also important drivers but the relationship is more complex (Section 6.2.1.3), so increasing temperature may lead to increased risk in some regions and decreased risk in others (IPCC, 2012). Mitigation and geoengineering strategies address the risk of anthropogenic climate change, and therefore the risk of hazards driven by temperature increases. A detailed synthesis report is given by the IPCC (2007c).

6.3.1.1 Mitigation

Mitigation strategies aim to reduce anthropogenic climate change by reducing the positive (warming) anthropogenic forcings. Mitigation is an extensive topic, global in reach (because GHGs have global effects) and politically charged (for example, developed nations are largely responsible for past GHG emissions but developing nations are most vulnerable to impacts). A few of the issues are outlined here; a more detailed discussion is given by the IPCC (2007c).

Most strategies for mitigating anthropogenic climate change focus on the sources of anthropogenic GHG emissions, by reducing usage of GHG-emitting processes (e.g. increased insulation of buildings, reduced air travel or reduced livestock consumption) or reducing their GHG intensity (e.g. energy efficiency, nuclear or renewable energy and fuel, carbon capture and storage, or low-methane agriculture). Some strategies focus on the sinks of GHGs, reducing atmospheric GHG concentrations by increasing uptake to terrestrial (e.g. trees, soil litter) or ocean reservoirs. Inertia in the climate system (due to long-timescale processes) and in societal change mean that most mitigation strategies are limited in the effect they can have on short-term (annual to decadal) climate change.

Some mitigation strategies have been put in place at the global scale. An international treaty was adopted in 1992 (United Nations Framework Convention on Climate Change, UNFCCC: http://unfccc.int), with legally binding GHG emissions targets agreed in 1997

and annual meetings to assess progress (IPCC, 2007c). National strategies of varying strength have also been put in place; the UK is currently unique in introducing a long-term legally binding framework, the Climate Change Act of 2008 (details can be found at http://www.legislation.gov.uk).

One of the controversial aspects of mitigation is the cost of large-scale societal and economic change motivated by predictions that are (inevitably) uncertain; for example, regional extreme temperature changes are poorly constrained by global warming targets (Clark *et al*., 2010). The UNFCCC explicitly applies the precautionary principle: 'where there are threats of serious or irreversible damage, lack of full scientific certainty should not be used as a reason for postponing such measures' (IPCC, 2007c: Article 3.3:). However, it is difficult to ascertain the most effective courses of action, and the best time frame for their implementation. A well-known attempt at quantifying the relative costs of action and inaction is the *Stern Review on the Economics of Climate Change*, an extensive report prepared by economists for the UK government, which estimated that if no action were taken the overall costs and risks of climate change would be equivalent to losing 5–20% or more of global GDP each year and that spending around 1% of global GDP each year to reduce GHG emissions would avoid the worst impacts (Stern, 2007). The assumptions and methodology of the *Stern Review* have been criticised (e.g. Weitzman, 2007): in particular, the use of a low discount rate (reduction in future costs of taking action).

6.3.1.2 Geoengineering

Geoengineering is deliberately changing the climate (Lenton and Vaughan, 2009; Irvine and Ridgwell, 2009). It has not yet been undertaken, except in small-scale studies (e.g. Boyd *et al*., 2000). Some proposed strategies aim to reduce the positive (warming) forcings, while others increase the negative (cooling) forcings. These correspond to the two possible methods of changing the earth's energy balance: increasing the outgoing (longwave) radiation, or reducing the incoming solar (shortwave) radiation.

The former methods decrease the forcing due to GHGs by extracting them from the atmosphere with technological (e.g. artificial trees: Lackner, 2009) or biological (e.g. fertilisation of ocean phytoplankton: Aumont and Bopp, 2006; afforestation, charcoal and wood storage) methods. These can also be considered mitigation strategies.

The latter type, known as solar radiation management (SRM), increase planetary reflectivity (albedo). Currently the most discussed method is injection of particles into the air to increase the reflectivity of the atmosphere or clouds (sulphate aerosols: e.g. Crutzen, 2006; water droplets: e.g. Ricke *et al*., 2010); other suggestions include increasing the reflectivity of urban structures (Oleson *et al*., 2010) or crops (Ridgwell *et al*., 2009), or positioning reflective structures between the earth and sun (Angel, 2006). One advantage of SRM over GHG reduction is that it could potentially cool the climate much more quickly. Different SRM methods have different strengths and weaknesses: for example, structures in space are thought to be prohibitively expensive, while the reversibility of particle injection methods (due to their short lifetime in the atmosphere) is an advantage if unforeseen negative impacts were to occur.

Risk management with SRM has several disadvantages relative to GHG reduction. It might allow anthropogenic GHG emissions to continue at their present rate, avoiding large-scale societal and economic changes, but would need to be maintained indefinitely: if it were stopped, rapid warming from the underlying GHG forcing would occur. SRM does not address carbon dioxide dissolution in the ocean, which may have significant negative impacts on calcifying marine organisms (Turley *et al*., 2010). Furthermore, SRM does not exactly cancel GHG forcing, because the earth system responds differently to different types of forcing (IPCC, 2007a). While global mean temperature could be returned to pre-industrial levels, this is unlikely to be the case for regional precipitation, drought, and storms (e.g. Ricke *et al*., 2010).

The global impacts of geoengineering strategies are explored with climate models. There are therefore large uncertainties in their effectiveness and the likelihood of negative impacts, particularly for methods that alter complex, nonlinear parts of the earth system such as the balance of ocean ecosystems or the hydrological cycle (IPCC, 2007c). A multi-model intercomparison project to evaluate the effects of stratospheric geoengineering with sulphate aerosols is currently underway (Kravitz *et al*., 2011). There are also enormous challenges in implementation, both practical and legal: for example, potentially setting the climatic needs and the rights of societies in different parts of the world against each other. It is therefore currently considered a method of last resort, or at least a method of support of, rather than a replacement for, mitigation.

6.3.2 Adaptation

Adaptation is management of risk by improving the ability to live with climate change: increasing resilience to minimise the severity of the impacts. Some categories of adaptation (Wilby *et al*., 2009) include resource management (e.g. of reservoirs), new infrastructure (e.g. flood defences) and behavioural change (e.g. diversification and genetic engineering of crops: CCSP, 2008; IPCC, 2007b). A detailed synthesis report is given by the IPCC (2007b).

Adapting to the impacts of extreme weather is a continuous process through human history (Chapter 5; IPCC, 2007b), but in the future these events may move outside the bounds of past experience in severity, frequency and location, requiring greater action and perhaps 'transformational' changes (IPCC, 2012). Measures to adapt to hazards in a changing climate are taken either as a consequence of past changes (e.g. increased use of artificial snow-making by the ski industry: IPCC, 2007b) or projections of future change. But the latter are uncertain, sometimes contradictory (Wilby and Dessai, 2010), and adaptation measures expensive, so strategies motivated by projections are generally 'no-' or 'low-regrets' options ('climate-influenced' decisions: Willows and Connell, 2003) that address multiple types of hazard, present and future risks and societal benefits such as water resource planning, coastal defence, social welfare, quality of life and livelihoods (IPCC, 2007b; Wilby and Dessai, 2010; IPCC, 2012). This helps to avoid competition for resources between adaptation, mitigation and disaster risk management (IPCC, 2012). Public pressure

for adaptation measures might increase with the occurrence of extreme weather events (IPCC, 2007c), though Hulme *et al.* (2011) argue that attempts to attribute extreme weather events to anthropogenic causes (e.g. Stott *et al.*, 2004; Min *et al.*, 2011; Pall *et al.*, 2011) are likely to 'increase the political and ethical complexities of these decisions'. Decisions are inevitably made with multiple objectives, which may be conflicting: 'maladaptations' are those that constrain adaptation by increasing exposure or vulnerability (e.g. developing housing in vulnerable areas: Willows and Connell, 2003; Chapter 5).

Adaptation has the potential to reduce future impacts even if the frequency or severity of hazards increase: for example, increasing temperatures and amplification of the hydrological cycle may increase the frequency and severity of droughts, but improvements in managing water resources mean this does not necessarily lead to more frequent restrictions in supply. The most successful adaptation strategies are iterative, with continuing dialogue between scientists and risk managers to improve the transfer, interpretation and usefulness of risk assessment information. Unfortunately, the cascading chain of uncertainties mean the use of climate scenarios for adaptation is 'lagging [their use in impact assessment] by nearly a decade' (Wilby *et al.*, 2009); the authors review current uncertainties in climate risk within an adaptation context.

Many risk management frameworks for climate change have been proposed (e.g. IPCC, 1994; Willows and Connell, 2003; Johnson and Weaver, 2009; IPCC, 2012). The 'risk, uncertainty and decision-making framework' (Figure 6.1; Willows and Connell, 2003) of the UK Climate Impacts Programme (http://www.ukcip.org.uk) is intended for users with some knowledge of climate risks that want to fully understand these risks and obtain a good understanding of adaptation options. The framework is described in a comprehensive technical document with detailed guidance, methods and options for implementation (Willows and Connell, 2003). Additional frameworks and tools are provided, based on the same underlying principles but aimed at different users, including an 'Adaptation Wizard' for users new to adaptation. Detailed case studies are given: for example, identifying the potential increasing risks of extreme weather for a UK container port (e.g. power supply disruption and damage, crane and pilot stoppages, port closures). The UKCIP framework has been used extensively by a wide variety of stakeholders to assess impacts and risk. A qualitative risk analysis with this framework can be followed with a quantitative analysis using the most recent UK Climate Projections (UKCP09: Murphy *et al.*, 2009), which contain more information and exploration of uncertainty than the previous projections (UKCIP02: Hulme *et al.*, 2002) and are expressed in probabilistic terms.

6.3.3 Communication

There are a number of problems in communicating the risks of future climate change and extreme weather hazards. A broad overview is given by Hulme (2009). The public receive conflicting messages from the media about whether the climate will change significantly (in contrast with the broad consensus within the scientific community on the large-scale
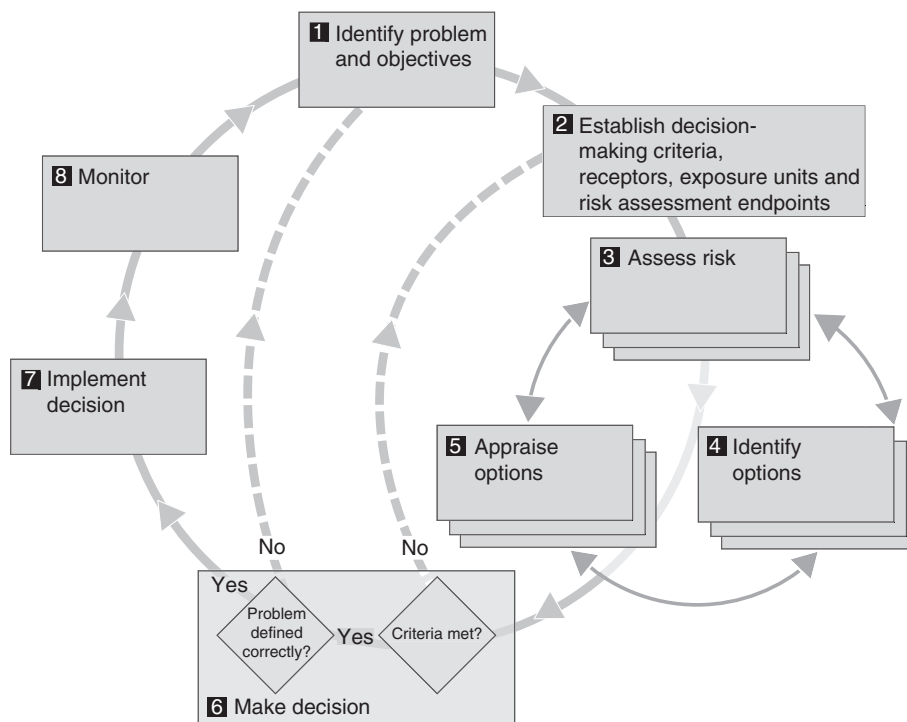
Figure 6.1 The UKCIP risk, uncertainty and decision-making framework (Willows and Connell, 2003 © UKCIP, 2011).

aspects: IPCC, 2007a) and whether individual extreme weather events are attributable to human causes (a reflection of this rapidly changing research area: e.g. Min *et al.*, 2011; Pall *et al.*, 2011). Communication of climate change risks is a wide-ranging and developing area. Some issues are outlined here.

Many difficulties arise from mental models and cognitive biases, where differences in framing, level of detail, personal and cultural experiences and narrative affect the interpretation of risk (Patt and Schrag, 2003; Marx *et al.*, 2007; Mastrandrea *et al.*, 2010; Morgan and Mellon, 2011; Sterman, 2011; Spiegelhalter *et al.*, 2011). Sterman (2011) describes some of the difficulties non-experts can have in understanding climate-related risks, including the long timescales, feedbacks and the complexity of the system, and makes a number of policy recommendations including a wider use of interactive simulations for learning; simple climate models with graphical user interfaces (e.g. Price *et al.*, 2009) could be used for this purpose.

Scientific literacy plays a part, but even well-informed users may be confused by the balance of consensus and uncertainty in the IPCC Assessment Reports (Socolow, 2011). These reports have guidelines that standardise and, over the years, refine the communication of uncertainty with calibrated language to reflect probabilistic predictions. In the Fourth Assessment Report (IPCC, 2007a) these included categories such as 'likely', corresponding to 60–90%

probability. In the *Special Report on Managing the Risks of Extreme Events and Disasters* (IPCC, 2012) and the Fifth Assessment Report, the calibrated intervals are cumulative rather than exclusive, so that 'likely' corresponds to 66–100% probability, and additional language has been added to reflect confidence of experts in the validity of a finding (Mastrandrea *et al.*, 2010, 2011). The associated quantitative intervals for each word or phrase are given at the end of the IPCC reports and not, in general, repeated throughout or in dissemination of material to the public. The danger in using calibrated language is that these words are interpreted differently by different people (Wallsten *et al.*, 1986; Patt and Schrag, 2003; Wardekker *et al.*, 2008; Morgan and Mellon, 2011). In fact, climate scientists also interpret and use the guidelines in contradictory ways (Spiegelhalter and Riesch, 2011; Mastrandrea and Mach, 2011).

There is no 'one size fits all' approach to communication or visualisation of risk (Chapter 5; IPCC, 2012). Spiegelhalter *et al.* (2011) give some helpful guidelines, including the need for multiple formats and the importance of iteration with user involvement. The most recent projections by the UK Climate Impacts Programme are available online (http://ukclimateprojections.defra.gov.uk) in the form of pre-prepared maps and graphs and an interactive tool for customised output. A separate map is provided for each combination of several weather variables (e.g. annual mean maximum temperature, winter precipitation), several locations (UK administrative regions), three emissions scenarios (low, medium, high), five probability thresholds (10%, 33%, 50%, 67%, 90%) and three future periods (2020s, 2050s, 2080s). Figure 6.2 shows maps for three probability
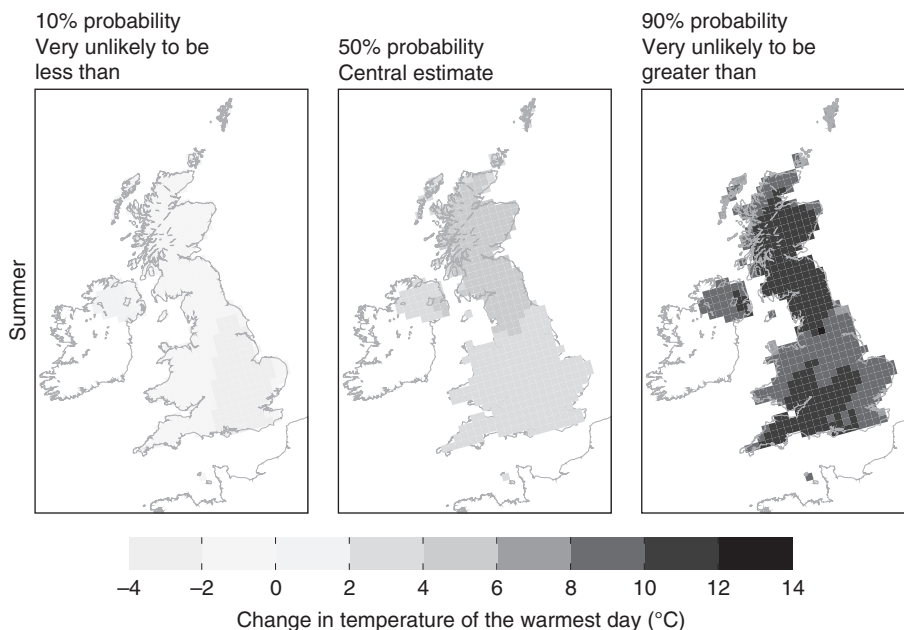

Figure 6.2 Maps from the UK Climate Impacts Programme showing 10%, 50% and 90% probability levels of changes to the temperature of the warmest day of the summer, by the 2080s, under the medium-emissions scenario (Murphy *et al.*, 2009 © UK Climate Projections, 2009).

thresholds for the change in temperature of the warmest day of the summer, by the 2080s, for the medium-emissions scenario. Explanatory phrases such as 'central estimate' and 'very unlikely to be greater than' are included to clarify the meaning of the probability levels. The separation of the different variables and scenarios makes large amounts of forecast information available, even without the customisation tool. These are used extensively by decision-makers such as local government agencies. However, there is little empirical evidence on the effectiveness of visualisation methods for climate change risks (Pidgeon and Fischhoff, 2011; Stephens *et al*., 2012).

6.4 Summary

Risk assessment and risk management for extreme weather in a changing climate is essential for global society, but it has many challenges. The limited predictability of future anthropogenic forcings is addressed by making predictions for a set of plausible storylines and scenarios, but these might not cover the range of possible futures. The effects of feedbacks are difficult to simulate for some important parts of the earth system, such as clouds and the carbon cycle, which means long-term projections are subject to substantial uncertainties. Short-term decadal predictions are highly desirable for policy-making and planning; methods are being learned from numerical weather prediction, but forecasts are dominated by internal variability and this remains one of the most challenging research areas. Extremes of climate are even more difficult to simulate than the mean, because observations are rarer (for model calibration and validation) and state-of-the-art climate models too expensive to run multi-centennial simulations. The most difficult weather hazards to simulate are extreme precipitation and tropical cyclones, because they are driven by processes on small spatial scales, and drought, which is defined by multiple variables and indices.

Physical modelling is crucial to incorporate the effects of changing boundary conditions. A spectrum of physical climate models exists, varying from complex, high-resolution regional models to simpler and lower-resolution global models. There is a trade-off, given finite computational resources, between climate model complexity (to reduce biases) and the length and number of simulations (to simulate extremes and quantify uncertainty). Physical impacts models (e.g. the effect of extreme weather on health and the environment) exist, but most impacts models are statistical and implicitly assume the hazard–impact relationship does not change through time.

Quantification of uncertainty incorporates all the challenges described in Chapter 5 and more. In particular, the multi-decadal timescale on which climate is defined rules out traditional validation methods used in numerical weather prediction. Models are mainly evaluated with hindcasts rather than forecasts, and repeated tests are made across spatial locations to partially compensate for the lack of repeated tests in time. Ensembles explore uncertainty in climate forcing, initial conditions, model parameters, and model structure. The significant computational expense of climate models usually limits the size of these ensembles to dozens, but simpler models (e.g. pattern scaling) and statistical models

(e.g. emulation) can supplement their predictions, and distributed computing has provided the resources to generate very large numbers of simulations. Not all ensembles are created equally: parametric and initial condition uncertainty (the former more important in centennial predictions, the latter in decadal) may be explored systematically and integrated out for a single model, but a suite of predictions for different scenarios or with different climate models must be treated individually or at least, in the case of multiple models, handled with care. It is essential, but very difficult, to estimate model structural uncertainty, because it represents the degree of realism of a climate model in an extremely high dimensional and correlated space; expert elicitation and multi-model ensembles are currently the main contenders. Model validation under uncertainty is further complicated by the low availability of observations. Reconstructions of past climates can help with this, though these require estimates of uncertainties to be useful. The causal chain under consideration is long – from emissions to concentrations, global climate, regional climate, extremes, impacts and, finally, risk management – and this is reflected in the relative uptake of methods to quantify and propagate uncertainty at each stage. It is rare for studies to propagate more than about three of these sources of uncertainty in any one assessment.

Potential risk management strategies are mitigation and geoengineering (reducing the human warming impacts on climate, or increasing the cooling impacts), adaptation (increasing resilience to extreme weather hazards today and in a changing climate) and communication (increasing awareness of the changing risks). Mitigation and geoengineering strategies are global in their effects, so they require international or global agreement to implement. The relative efficiency of mitigation, geoengineering and adaptation strategies is assessed with projections for different plausible future scenarios. However, the complex feedbacks and long causal chains of interactions between parts of the earth system and driving forces such as humans and the sun introduce many cascading uncertainties. This can lead to vague or contradictory risk assessments that are difficult to use in decision-making. To address this, risk management searches for 'no regrets' options that are flexible and/or yield additional benefits such as reduced costs (e.g. increased energy efficiency) or societal benefits (e.g. increased food security). Communication of risk is extremely challenging, given the long timescale, complex and statistical nature of climate and uncertainty in predictions. There have been improvements in the use of calibrated language, but little evaluation of the success of different visualisation methods.

The earth's climate is at once familiar and mysterious, and the weather hazards that occur in its extremes are among the most formidable aspects and the most challenging to predict. This chapter has outlined current methods and challenges of simulating the effect of changing boundary conditions on extreme weather, the statistical modelling that supports it and the types of risk management under consideration. The area cannot be done justice here, and is changing very rapidly, so the reader is encouraged to explore the literature further.

Acknowledgements

References

Allen, C., Macalady, A., Chenchouni, H., *et al.* (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* **259** (4): 660–684.

Alo, C. A. and Wang, Guiling (2008) Hydrological impact of the potential future vegetation response to climate changes projected by 8 GCMs. *Journal of Geophysical Research – Biogeosciences* **113** (G3): G03011.

Andronova, N. and Schlesinger, M. (2001) Objective estimation of the probability density function for climate sensitivity. *Journal of Geophysical Research* **106** (22): 605–622.

Angel, R. (2006) Feasibility of cooling the Earth with a cloud of small spacecraft near the inner Lagrange point (L1). *Proceedings of the National Academy of Sciences* **103** (46): 17184–17189.

Annan, J. D. and Hargreaves, J. C. (2010) Reliability of the CMIP3 ensemble. *Geophysical Research Letters* **37**: L02703.

Aumont, O. and Bopp, L. (2006) Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochemical Cycles* **20** (2): GB2017.

Barnett, D., Brown, S., Murphy, J., *et al.* (2006) Quantifying uncertainty in changes in extreme event frequency in response to doubled CO2 using a large ensemble of GCM simulations. *Climate Dynamics* **26** (5): 489–511.

Benestad, R. E., Hanssen-Bauer, I. and Chen, D. (2008) *Empirical-Statistical Downscaling*, Hackensack, NJ: World Scientific Publishing Co. Pte. Ltd.

Beniston, M., Stephenson, D., Christensen, O., *et al.* (2007) Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change* **81**: 71–95.

Best, M. J., Pryor, M., Clark, D., *et al.* (2011) The Joint UK Land Environment Simulator (JULES) model description: Part 1 – Energy and water fluxes. *Geoscientific Model Development* **4** (3): 677–699.

Betts, R. A., Boucher, O., Collins, M., *et al.* (2007) Projected increase in continental runoff due to plant responses to increasing carbon dioxide. *Nature* **448** (7157): 1037–1041.

Bjarnadottir, S., Li, Y. and Stewart, M. (2011) A probabilistic-based framework for impact and adaptation assessment of climate change on hurricane damage risks and costs. *Structural Safety* **33** (3): 173–185.

Bony, S. and Dufresne, J. (2005) Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters* **32** (20): L20806.

Boyd, P. W., Watson, A., Law, C., *et al.* (2000) A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature* **407** (6805): 695–702.

Braconnot, P., Otto-Bliesner, O., Harrison, S., *et al*. (2007) Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum: Part 1 – experiments and large-scale features. *Climate of the Past* **3** (2): 261–277.

Brown, S. J., Caesar, J. and Ferro, C. A. T. (2008) Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research – Atmospheres* **113** (D5): D05115.

Buontempo, C., Lørup, J. and Antar, M. (2009) Assessing the impacts of climate change on the water resources in the Nile Basin using a regional climate model ensemble. *Earth and Environmental Science* **6**: 292017.

Burke, E. J. and Brown, S. J. (2008) Evaluating uncertainties in the projection of future drought. *Journal of Hydrometeorology* **9** (2): 292–299.

Burke, E. J., Perry, R. H. J. and Brown, S. J. (2010) An extreme value analysis of UK drought and projections of change in the future. *Journal of Hydrology* **388** (1–2): 131–143.

CCSP (2008) *Climate models: An Assessment of Strengths and Limitations, Department of Energy*, Washington, DC: Office of Biological and Environmental Research.

Challinor, A., Wheeler, T., Craufurd, P., *et al*. (2004) Design and optimisation of a large-area process-based model for annual crops. *Agricultural and Forest Meteorology* **124**: 99–120.

Challinor, A., Wheeler, T., Slingo, J., *et al*. (2005) Quantification of physical and biological uncertainty in the simulation of the yield of a tropical crop using present-day and doubled CO2 climates. *Philosophical Transactions of the Royal Society B – Biological Sciences* **360** (1463): 2085–2094.

Challinor, A. J., Wheeler, T., Hemming, D., *et al*. (2009) Ensemble yield simulations: crop and climate uncertainties, sensitivity to temperature and genotypic adaptation to climate change. *Climate Research* **38** (2): 117–127.

Christensen, J. H., Carter, T. R. and Giorgi F. (2002) PRUDENCE employs new methods to assess European climate change. *EOS* **83**: 13.

Clark, D. B., Mercado, L., Sitch, S., *et al*. (2011) The Joint UK Land Environment Simulator (JULES) model description: Part 2 – carbon fluxes and vegetation dynamics. *Geoscientific Model Development* **4** (3): 701–722.

Clark, R. T., Brown, S. J. and Murphy, J. M. (2006) Modeling northern hemisphere summer heat extreme changes and their uncertainties using a physics ensemble of climate sensitivity experiments. *Journal of Climate* **19** (17): 4418–4435.

Clark, R. T., Murphy, J. M. and Brown, S. J. (2010) Do global warming targets limit heatwave risk? *Geophysical Research Letters* **37**: L17703.

Clune, T. L. and Rood, R. B. (2011) Software testing and verification in climate model development. *IEEE Software* **28** (6): 49–55.

Collins, M., Booth, B., Bhaskaran, B., *et al*. (2011) Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics* **36**: 1737–1766.

Cooley, D. (2009) Extreme value analysis and the study of climate change. *Climatic Change* **97**: 77–83.

Cox, P., Betts, R., Jones, C., *et al*. (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408** (6809): 184–187.

Crucifix, M. (2006) Does the Last Glacial Maximum constrain climate sensitivity? *Geophysical Research Letters* **33** (18): L18701.

Crutzen, P. J. (2006) Albedo enhancement by stratospheric sulfur injections: a contribution to resolve a policy dilemma? *Climatic Change* **77** (3–4): 211–220.

Dai, A. (2011) Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climatic Change* **2**: 45–65.

Diffenbaugh, N., Pal, J., Trapp, R., *et al*. (2005) Fine-scale processes regulate the response of extreme events to global climate change. *Proceedings of the National Academy of Sciences of the United States of America* **102** (44): 15774–15778.

Dlugolecki A., *et al*. (2009) Climate change and its implications for catastrophe modelling. In *Coping with Climate Change: Risks and Opportunites for Insurers*, London: Chartered Insurance Institute, pp. 1–23.

Eames, M., Kershaw, T. and Coley, D. (2011) On the creation of future probabilistic design weather years from UKCP09. *Building Services Engineering Research & Technology* **32** (2): 127–142.

Easterbrook, S. M. and Johns, T. C. (2009) Engineering the software for understanding climate change. *Computing in Science & Engineering* **11** (6): 64–74.

Edwards, T. L., Crucifix, M. and Harrison, S. P. (2007) Using the past to constrain the future: how the palaeorecord can improve estimates of global warming. *Progress in Physical Geography* **31** (5): 481–500.

Emanuel, K., Sundararajan, R. and Williams, J. (2008) Hurricanes and global warming: results from downscaling IPCC AR4 simulations. *Bulletin of the American Meteorological Society* **89** (3): 347–367.

Fildes, R. and Kourentzes, N. (2011) Validation and forecasting accuracy in models of climate change. *International Journal of Forecasting* **27** (4): 968–995.

Fischer, E. M. and Schär, C. (2008) Future changes in daily summer temperature variability: driving processes and role for temperature extremes. *Climate Dynamics* **33** (7–8): 917–935.

Fischer, E. M., Lawrence, D. M. and Sanderson, B. M. (2010) Quantifying uncertainties in projections of extremes: a perturbed land surface parameter experiment. *Climate Dynamics* **37**: 1381–1398.

Foley, A. M. (2010) Uncertainty in regional climate modelling: a review. *Progress in Physical Geography* **34** (5): 647–670.

Fowler, H. J., Cooley, D., Sain, S., *et al*. (2010) Detecting change in UK extreme precipitation using results from the climateprediction.net BBC climate change experiment. *Extremes* **13** (2): 241–267.

Fox-Rabinovitz, M., Cote, J., Dugas, B., *et al*. (2008) Stretched-grid Model Intercomparison Project: decadal regional climate simulations with enhanced variable and uniform-resolution GCMs. *Meteorology and Atmospheric Physics* **100** (1–4): 159–178.

Furrer, E. M. and Katz, R. W. (2008) Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resources Research* **44** (12): W12439.

Giorgi, F. and Mearns, L. (2002) Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the 'reliability ensemble averaging' (REA) method. *Journal of Climate* **15** (10): 1141–1158.

Goldstein, M. and Rougier, J. C. (2004) Probabilistic formulations for transferring inferences from mathematical models to physical systems. *SIAM Journal on Scientific Computing* **26** (2): 467–487.

Gosling, S. N., Lowe, J., McGregor, G., *et al*. (2009) Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. *Climatic Change* **92** (3–4): 299–341.

Gosling, S. N., McGregor, G. R. and Lowe, J. A. (2011) The benefits of quantifying climate model uncertainty in climate change impacts assessment: an example with heat-related mortality change estimates. *Climatic Change* **112** (2): 1–15.

Gregoire, L., Valdes, P., Payne, A., *et al*. (2010) Optimal tuning of a GCM using modern and glacial constraints. *Climate Dynamics* **37** (3–4): 1–15.

Hanel, M. and Buishand, T. A. (2011) Analysis of precipitation extremes in an ensemble of transient regional climate model simulations for the Rhine basin. *Climate Dynamics* **36**: 1135–1153.

Hanel, M., Buishand, T. A. and Ferro, C. A. T. (2009) A nonstationary index flood model for precipitation extremes in transient regional climate model simulations. *Journal of Geophysical Research – Atmospheres* **114**: D15107.

Hansen, J., Fung, A., Lacis, D., *et al*. (1988) Global Climate changes as forecast by Goddard Institute for Space Studies 3-Dimensional model. *Journal of Geophysical Research – Atmospheres* **93**: 9341–9364.

Hargreaves, J. C. (2010) Skill and uncertainty in climate models. *Wiley Interdisciplinary Reviews: Climatic Change* **1** (4): 556–564.

Harris, G. R., Sexton, D., Booth, B., *et al*. (2006) Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations. *Climate Dynamics* **27** (4): 357–375.

Hawkins, E. (2011) Our evolving climate: communicating the effects of climate variability. *Weather* **66** (7): 175–179.

Hawkins, E. and Sutton, R. (2009) The Potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society* **90** (8): 1095–1107.

Hemming, D., Buontempo, C., Burke, E., *et al*. (2010) How uncertain are climate model projections of water availability indicators across the Middle East? *Philosophical Transactions of the Royal Society A – Mathematical, Physical and Engineering Sciences* **368** (1931): 5117–5135.

Hewitson, B. (2003) Developing perturbations for climate change impact assessments. *Eos* **84** (35): 337–348.

Ho, C., Stephenson, D., Collins, M., *et al*. (2012) Calibration strategies: a source of additional uncertainty in climate change projections. *Bulletin of the American Meteorological Society* **93** (1): 21–26.

Hulme, M., O'Neill, S. J. and Dessai, S. (2011) Is weather event attribution necessary for adaptation funding? *Science* **334** (6057): 764–765.

Hulme, M. (2009) *Why We Disagree About Climatic Change*, Cambridge: Cambridge University Press.

Hulme, M., Jenkins, G. J., Lu, X. *et al*. (2002) Climate Change Scenarios for the United Kingdom: The UKCIP02 Scientific Report, Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK. 120pp.

IPCC (1994) *Technical Guidelines for Assessing Climate Change Impacts and Adaptations*. London: University College London and the Center for Global Environmental Research.

IPCC (2007a) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.

IPCC (2007b) *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.

IPCC (2007c) *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.

IPCC (2012) Summary for Policymakers. In: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (C. B. Field, V. Barros, T. F. Stocker, *et al*. (eds.)). A Special Report of Working Groups I and II of the Intergovernmental Panel

on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 3–21.

Irvine, P. and Ridgwell, A. (2009) 'Geoengineering': taking control of our planet's climate. *Science Progress* **92** (2): 139–162.

Johnson, T. E. and Weaver, C. P. (2009) A framework for assessing climate change impacts on water and watershed systems. *Environmental Management* **43** (1): 118–134.

Jones, P. D., Kilsby, C., Harpham, C., *et al.* (2009) Projections of future daily climate for the UK from the Weather Generator. In *UK Climate Projections*, Exeter: Met Office Hadley Centre, pp. 1–48.

Katz, R. W. (2010) Statistics of extremes in climate change. *Climatic Change* **100** (1): 71–76.

Kennedy, M. and O'Hagan, A. (2001) Bayesian calibration of computer models. *Journal of the Royal Statistical Society Series B – Statistical Methodology* **63**: 425–450.

Kennedy, M. C., Anderson, C., Conti, S., *et al.* (2006) Case studies in Gaussian process modelling of computer codes. *Reliability Engineering and System Safety* **91** (10–11): 1301–1309.

Kharin, V. V., Zwiers, F., Zhang, X., *et al.* (2007) Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate* **20** (8): 1419–1444.

Kim, J.-H., Brown, S. J. and McDonald, R. E. (2010) Future changes in tropical cyclone genesis in fully dynamic ocean- and mixed layer ocean-coupled climate models: a low-resolution model study. *Climate Dynamics* **37** (3–4): 737–758.

Klocke, D., Pincus, R. and Quaas, J. (2011) On constraining estimates of climate sensitivity with present-day observations through model weighting. *Journal of Climate*, **24**: 6092–6099.

Knapp, A. K., Beier, C., Briske, D., *et al.* (2008) Consequences of more extreme precipitation regimes for terrestrial ecosystems. *Bioscience* **58** (9): 811–821.

Knutson, T., McBride, J., Chan, J., *et al.* (2010) Tropical cyclones and climate change. *Nature Geoscience* **3**: 157–163.

Knutti, R. (2008a) Should we believe model predictions of future climate change? *Philosophical Transactions of the Royal Society A – Mathematical, Physical and Engineering Sciences* **366** (1885): 4647–4664.

Knutti, R. (2008b) Why are climate models reproducing the observed global surface warming so well? *Geophysical Research Letters* **35** (18): L18704.

Knutti, R. (2010) The end of model democracy? *Climatic Change* **102**: 395–404.

Knutti, R., Stocker, T., Joos, F., *et al.* (2003) Probabilistic climate change projections using neural networks. *Climate Dynamics* **21** (3–4): 257–272.

Knutti, R., Furrer, R., Tebaldi, C., *et al.* (2010) Challenges in combining projections from multiple climate models. *Journal of Climate* **23** (10): 2739–2758.

Kravitz, B., Robock, A., Bouchers, O., *et al.* (2011) The Geoengineering Model Intercomparison Project (GeoMIP). *Atmospheric Science Letters* **12** (2): 162–167.

Kyselý, J. (2009) Recent severe heat waves in central Europe: how to view them in a long-term prospect? *International Journal of Climatology* **30**: 89–109.

Kyselý, J. and Dubrovský, M. (2005) Simulation of extreme temperature events by a stochastic weather generator: effects of interdiurnal and interannual variability reproduction. *International Journal of Climatology* **25** (2): 251–269.

Lackner, K. S. (2009) Capture of carbon dioxide from ambient air. *The European Physical Journal Special Topics* **176** (1): 93–106.

Lambert, S. J. and Boer, G. J. (2001) CMIP1 evaluation and intercomparison of coupled climate models. *Climate Dynamics* **17** (2–3): 83–106.

Leimbach, M., Bauer, N., Baumstark, L., *et al.* (2010) Mitigation costs in a globalized world: climate policy analysis with REMIND-R. *Environmental Modeling & Assessment* **15** (3): 155–173.

Lenton, T. M. and Vaughan, N. E. (2009) The radiative forcing potential of different climate geoengineering options. *Atmospheric Chemistry and Physics* **9** (15): 5539–5561.

Lorenz, E. (1975) Climatic predictability. In *The Physical Basis of Climate and Climate Modeling*. Geneva: World Meteorological Organization, pp. 132–136.

Lorenz, P. and Jacob, D. (2005) Influence of regional scale information on the global circulation: a two-way nesting climate simulation. *Geophysical Research Letters* **32** (18): L18706.

Lynn, B. H., Healy, R. and Druyan, L. M. (2008) Quantifying the sensitivity of simulated climate change to model configuration. *Climatic Change* **92** (3–4): 275–298.

Marx, S. M., Weber, E., Orlove, B., *et al.* (2007) Communication and mental processes: experiential and analytic processing of uncertain climate information. *Global Environmental Change* **17** (1): 47–58.

Maslin, M., Owen, M., Betts, R., *et al.* (2010) Gas hydrates: past and future geohazard? *Philosophical Transactions of the Royal Society A – Mathematical, Physical and Engineering Sciences* **368** (1919): 2369–2393.

Masson, V. (2006) Urban surface modeling and the meso-scale impact of cities. *Theoretical and Applied Climatology* **84**: 35–45.

Mastrandrea, M. D. and Mach, K. J. (2011) Treatment of uncertainties in IPCC Assessment Reports: past approaches and considerations for the Fifth Assessment Report. *Climatic Change* **108** (4): 659–673.

Mastrandrea, M. D., Fields, C. B., Stocker, T. F., *et al.* (2010) Guidance note for lead authors of the IPCC Fifth Assessment Report on consistent treatment of uncertainties. Intergovernmental Panel on Climate Change (IPCC). http://www.ipcc-wg2.gov/meetings/CGCs/index.html#UR.

McFarlane, N. (2011) Parameterizations: representing key processes in climate models without resolving them. *Wiley Interdisciplinary Reviews: Climatic Change* **2** (4): 482–497.

McGregor, G. (2011) Human biometeorology. *Progress in Physical Geography* **36** (1): 93–109.

McGuffie, K. and Henderson-Sellers, A. (2005) *A Climate Modelling Primer*, 3rd edn, Chichester: John Wiley and Sons.

Meehl, G. A. and Tebaldi, C. (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. *Science* **305**: 994–997.

Meehl, G. A., Arblaster, J. M. and Tebaldi, C. (2005) Understanding future patterns of increased precipitation intensity in climate model simulations. *Geophysical Research Letters* **32** (18): L18719.

Meehl, G. A., Cover, C., Delworth, T., *et al.* (2007) THE WCRP CMIP3 Multimodel dataset: a new era in climate change research. *Bulletin of the American Meteorological Society* **88** (9): 1383–1394.

Mendes, D. and Marengo, J. A. (2009) Temporal downscaling: a comparison between artificial neural network and autocorrelation techniques over the Amazon Basin in present and future climate change scenarios. *Theoretical and Applied Climatology* **100** (3–4): 1–9.

Min, S.-K., Zhang, X., Zwiers, F., *et al.* (2009) Signal detectability in extreme precipitation changes assessed from twentieth century climate simulations. *Climate Dynamics* **32** (1): 95–111.

Min, S., Zhang, X., Zwiers, F., *et al.* (2011) Human contribution to more-intense precipitation extremes. *Nature* **470** (7334): 378–381.

Mitchell, T. (2003) Pattern scaling: an examination of the accuracy of the technique for describing future climates. *Climatic Change* **60** (3): 217–242.

Morgan, M. G. and Mellon, C. (2011) Certainty, uncertainty, and climate change. *Climatic Change* **108** (4): 707–721.

Murphy, J. M., Sexton, D., Barnett, D., *et al*. (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430** (7001): 768–772.

Murphy, J. M., Booth, B., Collins, M., *et al*. (2007) A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philosophical Transactions of the Royal Society A – Mathematical, Physical and Engineering Sciences* **365** (1857): 1993–2028.

Murphy, J. M., Sexton, D., Jenkins, G., *et al*. (2009) Climate change projections. In *UK Climate Projections*. Exeter: Met Office Hadley Centre, pp. 1–194.

Murphy, J., Clark, R., Collins, M. *et al*. (2011) Perturbed parameter ensembles as a tool for sampling model uncertainties and making climate projections. *Proceedings of ECMWF Workshop on Model Uncertainty*, 20–24 June 2011, 209–220. http://www.ecmwf.int/newsevents/meetings/workshops/2011/Model_uncertainty/presentations/Murphy.pdf.

Nakićenović, N., *et al*. (2000) *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Naveau, P., Nogaj, M., Ammann, C., *et al*. (2005) Statistical methods for the analysis of climate extremes. *Comptes Rendus Geoscience* **337**: 1013–1022.

Nicolosi, V., Cancelliere, A. and Rossi, G. (2009) Reducing risk of shortages due to drought in water supply systems using genetic algorithms. *Irrigation and Drainage* **58** (2): 171–188.

O'Hagan, A. (2006) Bayesian analysis of computer code outputs: a tutorial. *Reliability Engineering and System Safety* **91**: 1290–1300.

Oleson, K. W., Bonan, G. B. and Feddema, J. (2010) Effects of white roofs on urban temperature in a global climate model. *Geophysical Research Letters* **37**: L03701.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994) Verification, validation, and confirmation of numerical models in the earth sciences. *Science* **263** (5147): 641–646.

Pall, P., Aina, T., Stone, D., *et al*. (2011) Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature* **470** (7334): 382–385.

Palmer, T. (2001) A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society* **127** (572): 279–304.

Palmer, T. and Räisänen, J. (2002) Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature* **415** (6871): 512–514.

Palmer, T. and Williams, P. (eds) (2009) *Stochastic Physics and Climate Modelling*, Cambridge: Cambridge University Press.

Palmer, T., Alessandri, A., Andersen, U., *et al*. (2004) Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society* **85** (6): 853–872.

Palmer, T., Shutts, G., Hegedorn, R., *et al*. (2005) Representing model uncertainty in weather and climate prediction. *Annual Review of Earth and Planetary Sciences* **33**: 163–193.

Palmer, T., Doblas-Reyes, F., Weisheimer, A., *et al*. (2008) Toward seamless prediction: calibration of climate change projections using seasonal forecasts. *Bulletin of the American Meteorological Society* **89** (4): 459–470.

Patt, A. and Schrag, D. (2003) Using specific language to describe risk and probability. *Climatic Change* **61**: 17–30.

Pennell, C. and Reichler, T. (2011) On the effective number of climate models. *Journal of Climate* **24** (9): 2358–2367.

Piani, C., Frame, D., Stainforth, D., *et al*. (2005) Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophysical Research Letters* **32** (23): L23825.

Pidgeon, N. and Fischhoff, B. (2011) The role of social and decision sciences in communicating uncertain climate risks. *Nature Climatic Change* **1** (1): 35–41.

Pirtle, Z., Meyer, R. and Hamilton, A. (2010) What does it mean when climate models agree? A case for assessing independence among general circulation models. *Environmental Science & Policy* **13** (5): 351–361.

Pope, V. and Davies, T. (2002) Testing and evaluating atmospheric climate models. *Computing in Science & Engineering* **4** (5): 64–69.

Price, A. R., Goswami, S., Johnson, M., *et al*. (2009) The Aladdin-2 launchpad: engaging the user community and broadening access to the GENIE model on local and remote computing resources. In UK e-Science All Hands Meeting 2009, Past Present and Future. Monday 7–9 December 2009, Oxford.

Qian, B., Gameda, S. and Hathoe, H. (2004) Performance of stochastic weather generators LARS-WG and AAFC-WG for reproducing daily extremes of diverse Canadian climates. *Climate Research* **26**: 175–191.

Räisänen, J. (2005) Impact of increasing CO_2 on monthly-to-annual precipitation extremes: analysis of the CMIP2 experiments. *Climate Dynamics* **24**: 309–323.

Räisänen, J. (2007) How reliable are climate models? *Tellus Series A – Dynamic Meteorology and Oceanography* **59** (1): 2–29.

Räisänen, J. and Palmer, T. (2001) A probability and decision-model analysis of a multimodel ensemble of climate change simulations. *Journal of Climate* **14** (15): 3212–3226.

Randall, D., Krueger, S., Bretherton, C., *et al*. (2003) Confronting models with data: the GEWEX cloud systems study. *Bulletin of the American Meteorological Society* **84** (4): 455–469.

Reichler, T. and Kim, J. (2008) How well do coupled models simulate today's climate? *Bulletin of the American Meteorological Society* **89** (3): 303–311.

Ricke, K. L., Morgan, M. G. and Allen, M. R. (2010) Regional climate response to solar-radiation management. *Nature Geoscience* **3** (8): 537–541.

Ridgwell, A., Singarayer, J., Hetherington, A., *et al*. (2009) Tackling regional climate change by leaf albedo bio-geoengineering. *Current Biology* **19** (2): 146–150.

Rodwell, M. J. and Palmer, T. N. (2007) Using numerical weather prediction to assess climate models. *Quarterly Journal of the Royal Meteorological Society* **133** (622): 129–146.

Rougier, J. (2007) Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change* **81** (3–4): 247–264.

Rougier, J. and Sexton, D. M. H. (2007) Inference in ensemble experiments. *Philosophical Transactions of the Royal Society A – Mathematical, Physical and Engineering Sciences* **365** (1857): 2133–2143.

Rougier, J., Sexton, D., Murphy, J., *et al*. (2009) Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments. *Journal of Climate* **22** (13): 3540–3557.

Roy, C. and Oberkampf, W. (2011) A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering* **200**: 2131–2144.

Ruddiman, W. (2003) The anthropogenic greenhouse era began thousands of years ago. *Climatic Change* **61** (3): 261–293.

Rummukainen, M. (2010) State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climatic Change* **1** (1): 82–96.

Sanderson, B. M. (2011) A multimodel study of parametric uncertainty in predictions of climate response to rising greenhouse gas concentrations. *Journal of Climate* **24** (5): 1362–1377.

Sanderson, B. M., Knutti, R., Aina, T., *et al*. (2008) Constraints on model response to greenhouse gas forcing and the role of subgrid-scale processes. *Journal of Climate* **21** (11): 2384–2400.

Scheffer, M., Brovkin, V. and Cox, P. M. (2006) Positive feedback between global warming and atmospheric CO2 concentration inferred from past climate change. *Geophysical Research Letters* **33** (10): L10702.

Schwierz, C., Kollner-Heck, P., Zenklusen Mutter, E., *et al*. (2010) Modelling European winter wind storm losses in current and future climate. *Climatic Change* **101**: 485–514.

Semenov, M. (2008) Simulation of extreme weather events by a stochastic weather generator. *Climate Research* **35**: 203–212.

Sexton, D. and Murphy, J. (2011) Multivariate probabilistic projections using imperfect climate models. Part II: robustness of methodological choices and consequences for climate sensitivity. *Climate Dynamics* **38**: 2543–2558.

Sexton, D., Murphy, J., Collins, M., *et al*. (2011) Multivariate probabilistic projections using imperfect climate models. Part I: outline of methodology. *Climate Dynamics* **38**: 2513–2542.

Smith, D. M., Cusack, S., Colman, A., *et al*. (2007) Improved surface temperature prediction for the coming decade from a global climate model. *Science* **317** (5839): 796–799.

Socolow, R. H. (2011) High-consequence outcomes and internal disagreements: tell us more, please. *Climatic Change* **108** (4): 775–790.

Soden, B. J. and Held, I. M. (2006) An assessment of climate feedbacks in coupled ocean–atmosphere models. *Journal of Climate* **19** (14): 3354–3360.

Spiegelhalter, D. J. and Riesch, H. (2011) Don't know, can't know: embracing deeper uncertainties when analysing risks. *Philosophical Transactions of the Royal Society A – Mathematical, Physical And Engineering Sciences* **369**: 4730–4750.

Spiegelhalter, D., Pearson, M. and Short, I. (2011) Visualizing uncertainty about the future. *Science* **333** (6048): 1393–1400.

Stainforth, D. A., Aina, T., Christensen, C., *et al*. (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* **433** (7024): 403–406.

Stensrud, D. J. (2009) *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*, Cambridge: Cambridge University Press.

Stephens, E. M., Edwards, T. L. and Demeritt, D. (2012) Communicating probabilistic information from climate model ensembles – lessons from numerical weather prediction. *Wiley Interdisciplinary Reviews: Climate Change* **3** (5): 409–426.

Sterl, A., Severijns, C., Dijkstra, H., *et al*. (2008) When can we expect extremely high surface temperatures. *Geophysical Research Letters* **35**: L14703.

Sterman, J. (2011) Communicating climate change risks in a skeptical world. *Climatic Change* **108**: 811–826.

Stern, N. H. (2007) *Stern Review on the Economics of Climate Change*, Cambridge: Cambridge University Press.

Stott, P. A. and Forest, C. E. (2007) Ensemble climate predictions using climate models and observational constraints. *Philosophical Transactions of the Royal Society A – Mathematical, Physical And Engineering Sciences* **365** (1857): 2029–2052.

Stott, P., Stone, D. A. and Allen, M. R. (2004) Human contribution to the European heatwave of 2003. *Nature* **432** (7017): 610–614.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A. (2011) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society* **93**: 485–498.

Tebaldi, C., Smith, R. L., Nychka, D., *et al*. (2005) Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multimodel ensembles. *Journal of Climate* **18** (10): 1524–1540.

Tebaldi, C., Hayhoe, K., Arblaster, J., *et al*. (2006) Going to the extremes. *Climatic Change* **79** (3–4): 185–211.

Torn, M. S. and Harte, J. (2006) Missing feedbacks, asymmetric uncertainties, and the underestimation of future warming. *Geophysical Research Letters* **33** (10): L10703.

Tsvetsinskaya, E., Mearns, L., Mavromatis, T., *et al*. (2003) The effect of spatial scale of climatic change scenarios on simulated maize, winter wheat, and rice production in the southeastern United States. *Climatic Change* **60**: 37–71.

Turley, C., Eby, M., Ridgwell, A., *et al*. (2010) The societal challenge of ocean acidification. *Marine Pollution Bulletin* **60** (6): 787–792.

Valdes, P. (2011) Built for stability. *Nature Geoscience* **4** (7): 414–416.

Vuuren, D. P. van, Edmonds, J., Kainuma, M., *et al*. (2011) The representative concentration pathways: an overview. *Climatic Change* **109**: 5–31.

Wang, G. L. (2005) Agricultural drought in a future climate: results from 15 global climate models participating in the IPCC 4th assessment. *Climate Dynamics* **25**: 739–753.

Wallsten, T. S., Zwick, R., Forsyth, B., *et al*. (1986) Measuring the vague meanings of probability terms. *Journal of Experimental Psychology – General* **115** (4): 348.

Wardekker, J. A., van der Sluijs, J. P., Kloprogge, P., *et al*. (2008) Uncertainty communication in environmental assessments: views from the Dutch science–policy interface. *Environmental Science & Policy* **11** (7): 627–641.

Webb, M. J., Senior, C., Sexton, D., *et al*. (2006) On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics* **27** (1): 17–38.

Weigel, A. P., Liniger, M. A., Appenzeller, C., *et al*. (2010) Risks of model weighting in multimodel climate projections. *Journal of Climate* **23** (15): 4175–4191.

Weisheimer, A. and Palmer, T. (2005) Changing frequency of occurrence of extreme seasonal temperatures under global warming. *Geophysical Research Letters* **32** (20): L20721.

Weitzman, M. L. (2007) A review of the Stern Review on the economics of climate change. *Journal of Economic Literature* **45** (3): 703–724.

Wilby, R. L. and Dessai, S. (2010) Robust adaptation to climate change. *Weather* **65** (7): 180–185.

Wilby, R., Dawson, C. and Barrow, E. (2002) SDSM: a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software* **17** (2): 147–159.

Wilby, R., Charles, S., Zorita, E., *et al*. (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods. Available from the DDC of IPCC TGCIA.

Wilby, R. L., Troni, J., Biot, Y., *et al*. (2009) A review of climate risk information for adaptation and development planning. *International Journal of Climatology* **29** (9): 1193–1215.

Wilks, D. S. and Wilby, R. L. (1999) The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23** (3): 329–357.

Willows, R. I. and Connell, R. K. (eds) (2003) *Climate Adaptation: Risk, Uncertainty and Decision-making*. Oxford: UKCIP.

Wright, G., Jack, L. and Swaffield, J. (2006) Investigation and numerical modelling of roof drainage systems under extreme events. *Building and Environment* **41** (2): 126–135.

Yang, Z. and Arritt, R. (2002) Tests of a perturbed physics ensemble approach for regional climate modeling. *Journal of Climate* **15** (20): 2881–2896.

Yokohata, T., Webb, M., Collins, M., *et al.* (2010) Structural similarities and differences in climate responses to CO2 increase between two perturbed physics ensembles. *Journal of Climate* **23** (6): 1392–1410.